GENOME EDITING WITH CRISPR-CAS9 IN THE ILLINOIS LONG TERM SELECTION EXPERIMENT

BY

STEPHEN JOSEPH JINGA

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Crop Sciences
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2019

Urbana, Illinois

Master's Committee:

Professor Stephen Moose, Chair
Assistant Professor Tiffany Jamann
Assistant Professor Sarah Hind

ABSTRACT

Recent advances in genome editing by Clustered Regularly Interspaced Palindromic Repeat (CRISPR) and CRISPR Associated Protein 9 (Cas9) have markedly increased our ability to characterize genes and use genetics to the benefit of agriculture. In this work, we utilize this technology to study the Illinois Long Term Selection Experiment (ILTSE), a unique germplasm resource for studies of genome evolution and genetic variants that contribute to phenotypic traits. ILTSE genotypes, Illinois High Protein 1 and Illinois Low Protein 1, create highly regenerable embryogenic type I callus, enabling transformation and genome editing approaches to characterize gene function. The *Lemon White 1 (Lw1)* locus was initially targeted as a proof of concept to generate albino plants easily detectable in a population of regenerated plants. Four guide RNAs were tested for their function using an *in vitro* Cas9 cleavage assay. CRISPR editing vectors were delivered to embryogenic calli using biolistics and transgenic events selected. Multiple albino plants indicative of biallelic mutations were recovered in the ILP1 genotypes at 1.5% efficiency; however none were produced from the IHP1 genotype. A second CRISPR experiment targeted the *L-Asparaginase (ASNase)* gene, which exhibits reduced gene expression in IHP1 compared to ILP1. The goal was to test whether reducing ASNase gene function can increase grain protein concentration in the ILP1 background. Four guide RNAs were designed and tested *in vitro* before delivery. Two ILP1 events were generated with novel *ASNase* deletion alleles. Limited T1 seed was recovered and will be used for future characterization.

ACKNOWLEDGMENTS

I am forever grateful to the numerous people I have the privilege to interact with at the University of Illinois. First, I would like to thank Professor Stephen Moose for the opportunity to conduct graduate level research in his lab and help him teach his undergraduate Biotechnology in Agriculture course. His knowledge and experience have been instrumental in helping me complete my research and advance my career in science and the study of plants. I also want to give thanks to all members of the Moose Lab, especially Brian Rhodes, Eddie Ross, and Jessica Bubert, for helping me progress through the program and for being great friends.

I have also had the privilege of working with numerous undergraduate students who have surpassed my expectations and will continue to thrive in their careers. Of these students, I would like to acknowledge Anna Parkinson for playing a large role in helping me keep up with the tissue culture and making sure things were on schedule. Other students including: Camilla Macias, Thomas Janas, Alex Bahnick, Grace Carlberg, Rob McCormick, Mara Swapp, and Dave Istanto have also contributed to helping me complete this work.

I would also like to thank my partner Jill Wallitschek for her continuous support throughout graduate school. She has contributed greatly to my growth in the sciences as well as a person. Finally, I would like to thank my parents John and Diane Jinga for encouraging me to always strive for excellence and always being there for me.

TABLE OF CONTENTS

CHAPTER 1: TISSUE CULTURE AND GENE EDITING OF THE *LEMON WHITE 1* LOCUS IN ILTSE LINES

INTRODUCTION

The Illinois Long Term Selection Experiment

The Illinois Long Term Selection Experiment (ITLSE) is the longest ongoing crop genetics experiment and is currently curated by the Moose Laboratory. The goals of this project, set forth by UIUC scientist Cyril Hopkins, were to change the chemical composition of corn kernels (Hopkins, 1899). This included creating varieties that had altered protein and oil levels to encompass the diverse uses for corn. In 1896, 163 ears of the open pollinated variety "Burr's White" were measured for protein and oil concentration. Ears that contained the lowest percentage protein were used to create the Illinois Low Protein (ILP) population and ears with the highest percentage created the Illinois High Protein (IHP) population. Similar principles created the Illinois High Oil (IHO) and Illinois Low Oil (ILO) populations. All populations were continued at the University of Illinois at Urbana-Champaign and subjected to directional selection based on the population goals. By cycle 10, it was reported that all populations had altered kernel composition consistent with their respective selected phenotypes (Smith, 1908). Continuing into the present day, the ILTSE has become a unique resource for studying economical traits in corn such as seed composition and nitrogen use. While normal kernel protein content ranges from 8-12% in maize, the IHP and ILP populations have a selection response that is greater than 20 standard deviations from the original population (Moose et al., 2004). In addition, other populations have been added to the experiment that reverses the direction of selection to find alleles that are most sensitive to selection for these traits.

The most recent cycles of the Illinois Protein Strains measure 30% kernel protein in IHP and 4% kernel protein in ILP (Figure 1). At cycle 90 of the experiment, inbred lines of the protein strains were created by multiple generations of selfing. When used as an inbred parent, the protein strains create a large difference in hybrid plant Nitrogen use (Uribelarrea et al., 2006). Storage proteins are the primary method of maize to store nitrogen, thus with the drastic differences in kernel protein content, these lines differ in many genes that contribute to overall kernel protein. Many QTL for kernel traits have been identified utilizing this germplasm. Goldman et al., (1993) reported 22 loci associated with protein concentration spread across all 10 chromosomes, some of which are suggested to be unique to the ILTSE due to agronomic differences from Burr's White to more popular varieties of corn. Thus, the ILTSE is unlike common varieties of corn grown today and may provide insights to studying selection, genome evolution, and the genes contributing to Nitrogen metabolism.

Maize tissue culture and ILTSE culture response

Maize tissue culture was first reported in 1975 by Green and Phillips. Articles published early in this field indicate that parental genotype, genotype by medium interactions, and genes that control somatic embryogenesis are major factors that allow plant regeneration (Willman et al., 1989; Wilkinson et al., 1987). There have been many studies done to assess which genotypes produce highly regenerable callus at high frequency, including Illinois High Oil, Illinois High Protein, Illinois Low Protein, and Illinois Reverse Low Protein (Duncan et al., 1985). The ability to generate this type of tissue enables

transformation. This ability to regenerate is not present in all maize lines, such as the sequenced inbred B73 and other public lines such as Mo17. Hi-II is a popular maize transformation line derived from the cross of the high regeneration line A188 with B73, and performs well due to its rapidly growing callus and high regeneration frequency (Armstrong et al., 1991). Despite Hi-II's performance as a transformation line, introducing genetic variation into a hybrid background can complicate subsequent functional genomics analyses. Direct transformation of inbred lines also offers advantages in further breeding efforts. Protocols for various inbred maize genotypes for transformation have been established and refined over the past few decades (Ishida et al., 2007, Raji et al., 2017)

There have been many methods developed that successfully integrate transgenes in maize, such as biolistics, electroporation, and *Agrobacterium tumefaciens* (Klein et al., 1987, D'Halluin et al., 1992, Songstad et al. 1996, Armstrong, 1999). The first described transgenic maize was developed in A188 protoplasts with a construct containing 35S:NPTII (Rhodes et al., 1988). This was an important pre-requisite for maize biotechnology, demonstrating that DNA from other kingdoms of life could function in a maize cell. These plants were not fertile however, and it wasn't until two years later when scientists at DeKalb produced the first fertile transgenic maize containing the PAT gene and the GUS reporter gene (Gordon-Kamm et al., 1990). This was the first step in a long cascade of events that led to the commercialization of transgenic corn and improved traits such as disease, insect, and herbicide resistance.

Compared to other crops, maize is considered recalcitrant to *Agrobacterium*-mediated transformation (Gordon-Kamm et al., 1990). This has been overcome by the use of an innovative particle bombardment system first outlined by Klein et al., 1987 that is still utilized today. Compared to *Agrobacterium*, biolistics has more flexibility for transforming multiple constructs simultaneously, thus improving efficiency and lowering genotype dependency (Raji et al., 2017). Despite the many successes of transformation using the methods mentioned above, overall the process is often time consuming and results in a limited number of successful transgenic events. One approach to overcome this obstacle was a recent breakthrough report that uses the morphogenic regulators *Baby Boom* and *Wuschel* to dramatically increase transformation frequency. In this study, the authors use these two genes to dramatically increase callus production in lines that were previously recalcitrant to traditional methods (Lowe et al., 2016, Jones et al., 2019).

Gene editing with CRISPR-Cas9

Recent advances in genome editing by Clustered Regularly Interspaced Short Palindromic Repeat (CRISPR) and CRISPR Associated Protein 9 (Cas9) have remarkably increased our ability to characterize genes by introducing mutations leading to loss-of-function alleles, manipulating expression, or introducing new sequences (Doudna et al., 2014). This system utilizes the Cas9 nuclease found naturally in bacterial immune systems and a programmable guide RNA (gRNA) comprised of crRNA and tracrRNA sequences that can be designed to target a specific genomic region (Jinek et al., 2012). Delivery of these components to a cell can generate double strand breaks (DSBs) in genomic DNA, directed by the 20 nt gRNA target sequence and protospacer adjacent

motif (PAM). Repair of DSBs by native cellular functions, primarily by error prone non-homologous end joining (NHEJ), results in random mutations at the target location in the genome. Site-specific mutations and insertions can result from homology directed DSB repair (HDR), which occurs at lower frequencies and requires a DNA template for correct resolution of the break.

Numerous reports have now demonstrated success with using CRISPR/Cas9 for targeted mutagenesis of the maize genome, for agronomic traits such as male sterility, improved drought tolerance, herbicide resistance and the waxy mutation (Liang et al., 2014; Svitashev et al., 2015; Shi et al., 2016, Svitashev et al., 2016, Xing et al., 2014). Mutations at specific sites generated by such methods are useful to study gene function, generate variation, and increase genetic gain by reducing the number of breeding cycles. Proof of concept gene editing in maize has been done by targeting the *Lemon White 1* (*Lw1*) gene to create albino plants (Feng et al., 2016, Feng et al., 2018). The easily observable mutant phenotype makes it an ideal locus to target when developing editing systems in new genotypes. *Lw1* is named as "Zm00001d033896" in version 4 of the B73 genome (Jiao et al., 2017) housed at MaizeGDB (www.maizegdb.org). This gene encodes for 4-hydroxy-3-methylbut-2-enyl diphosphate reductase and the protein functions in the non-mevalonate pathway that produces the isoprenoids isopentyl diphosphate (IPP) and dimethylallyl disphosphate (DMAPP). Loss of this enzyme impedes the plants ability to create both carotenoids and chlorophyll.

Multiplex Genome Editing

The use of multiple guide RNAs designed to target one or more genes is useful for generating knock-outs of multiple gene products, large deletions of a locus, and homologous recombination within a locus. Such experiments are termed as "multiplexing" and allow for greater flexibility within editing experiments. These strategies utilize polycistronic systems including: tRNA processing (Qi et al., 2016), self-cleaving ribozymes (Tang et al., 2016), and Csy4 ribonuclease from bacterial CRISPRs (Cermak et al., 2017). In maize, the tRNA processing system has been reported to express four guide RNAs in a single experiment (Qi et al., 2016). In contrast, prokaryotic CRISPRs utilize a complex of Csy proteins to process crRNAs expressed from a single transcript, in which multiple crRNAs are released for viral defense (Haurwitz et al., 2010). The Csy4 protein alone is sufficient for polycistronic gRNA processing, and this system has worked for multiplexing editing in other crops such as wheat, tomato, and alfalfa (Cermak et al., 2017). We demonstrate here the regeneration of the maize ILP1 inbred line with CRISPR/Cas9 edits induced by multiple gRNAs as a result of Csy4 processing.

MATERIALS AND METHODS

Construction of CRISPR-Csy4 vectors for biolistic transformation

Vectors used in these experiments were created using the plant genome engineering toolkit outlined in Čermák et al 2017 (Figure 4). Online resources and protocols using these systems are found on their website: http://cfans-pmorrell.oit.umn.edu/CRISPR_Multiplex/. Briefly, custom oligonucleotides for each

designed gRNA were ordered with corresponding *Sap*I and *Esp*3I sites to facilitate cloning into pMOD_B2103. Destination vectors and PCR products containing four unique gRNA-Csy4 cassette were ligated together in a Golden Gate reaction and resulting colonies screened for correct assembly by Sanger sequencing. This plasmid was then placed in a Golden Gate reaction using *Aar*I with plasmids pTRANS_120 for the 2x35S:NPTII backbone, pMOD_1510 containing ZmUbi:Csy4-TaCas9, and pMOD_C0000 empty HDR cassette. Clones were screened on LB media containing spectinomycin and Sanger sequenced around Golden Gate junctions for correct ligation.

***In vitro* digestion of gene targets with Cas9 Ribonucleoproteins**

In vitro Cas9 Ribonucleoprotein (RNP) digestions were performed as previously described in Liang et al., 2017 with minor changes. Cas9-6X HIS protein was isolated using a Ni-NTA purification outlined in Liang et al., 2017. Oligonucleotides for each designed guide RNA was annealed, phosphorylated with T4 Polynucleotide Kinase (NEB Catalog #M0201S), and cloned into pT7-gRNA (Addgene Plasmid #46759). Synthetic RNA molecules of each guide RNA target sequence and scaffold were created following NEB's T7 RNA Synthesis Kit (Catalog #E2040S). Purified PCR products encompassing all gRNA traget sites were amplified and eluted into RNAse free water. The Cas9-gRNA digestion reaction was set up using Cas9 (1 ug), gRNA (375 ng), PCR product (200 ng), and 10X Cas9 reaction buffer (20mM HEPES, pH 7.5, 150 mM KC, 10mM MgCl2, 0.5mM DTT). The reaction was incubated at 37°C for 1 hour for Cas9 digestion and 25°C for 15 minutes after addition of proteinase K and at 56°C for 10 minutes. Digested products were run on a 50 mL 2% TAE agarose gel at 90V for 90 min.

Plant Tissue Culture

Maize inbreds ILP1, IHP1, and H99 were grown in field plots at the University of Illinois at Urbana-Champaign Research and Education Center in Urbana, IL during the summer of 2018. Each plant was self-pollinated and harvested between 10-15 days after pollination when the embryo reached 1.5-2mm in length. Immature embryos were plated on N6 Dicamba (N6D) callus initiation medium in the dark and checked for callus growth 3-4 days after induction. Embryogenic callus growth was detected on embryos as a bumpy, popcorn-like appearance (Figure 2A, B). Calli were moved to N6 Osmotic (N6O) media four hours prior to bombardment. After transformation, each callus was placed on a selective N6 media containing 300mg/uL paramomycin and events were propagated over 8-12 weeks in dark conditions. Resistant calli were moved to a Regeneration I media containing 5mg/mL 6-BAP for four days and then transferred to media lacking 6-BAP for three additional days (Figure 1C). All putative events were moved into the light (16h on/ 8h off) to generate shoot tissues. Plantlets with developed shoot and root systems were moved to Regeneration II media (Figure 1D) and later acclimated to ambient humidity for greenhouse growth (Figure 1E). The full protocol for tissue culture is outlined in Appendix B.

Particle Bombardment

Protocols for particle bombardment were adapted from Frame et al., 2000. Briefly, 1.8mg 0.6μm gold particles were washed with 100% ethanol and sterile water. 600ng of plasmid DNA was mixed with gold particles and quickly mixed with 2.5M $CaCl_2$ and 0.1M sperimidine. The mixture was incubated on ice for two hours before being resuspended in

ethanol and added to macrocarriers (100ng DNA per shot) for bombardment. Each plate of embryos was shot two times.

Extraction of DNA from leaf tissue

DNA was extracted from regenerated plants following a modified CTAB protocol adapted from Porebski et al 1997. The samples were ground and resuspended in CTAB buffer containing RNase followed by incubation at 65°C for 30 minutes. A total of 600 µL of chloroform were added and mixed. The samples were centrifuged at 9000 rpm for 10 minutes and the top layer removed and pipetted into a new tube containing 450µL of room temperature isopropanol and mixed well. Samples were centrifuged at 12000 rpm for 10 minutes at 4°C, the supernatant decanted, and the DNA pellet washed with 70% ethanol. Finally, ethanol was removed and the DNA pellet resuspended in 100µL of water.

PCR amplification and sequencing of target sites

Genomic DNA was amplified using NEB's 5x Taq Master Mix (Cat #M0285L) using either the Lw1-F1 and Lw1-R2 or Lw1-F2 and Lw1-R5 primer pairs (Figure 5B). Amplicons were produced using the following protocol: Denaturation - 95°C for 30 seconds, Cycling (x35) – 95°C for 30 seconds, 55°C for 30 seconds, and 68°C for one minute, and a 68°C final extension step for 5 minutes. PCR products were run on a 1% agarose gel using Invitrogen's 1kB plus DNA ladder (Catalog # 10787018), visualized with ethidium bromide and UV light, and images recorded with GeneSnap software.

RESULTS

Reliable regeneration of Illinois Protein Strains

Maize transformation experiments are largely limited by genotypes that create highly regenerable embryogenic callus. Using methods adapted from Duncan et al., 1985 and protocols for type I callus (Iowa State Plant Transformation Facility) we initiated embryogenic callus cultures and regenerated fertile plants. Callus from IHP1 and ILP1 exhibit distinct morphogenic features 10 days in culture on N6D media in the dark at 25°C (Figure 2A, B). IHP1 callus grew at a faster rate and was more friable and soft compared to ILP1, which produced a rigid callus resembling type I embryogenic calli from other transformation lines such as H99. Regenerating the callus back into plantlets produced fertile corn plants that could be grown to seed in a greenhouse (Figure 2C, D, E).

Guide RNA design and validation *in vitro*

Three guide RNAs for *Lemon White 1 (Zm00001d033896)* were identified from version 4 the B73 reference genome (Jiao et al., 2017) using CHOPCHOP (http://chopchop.cbu.uib.no) and CRISPR-P 2.0 (http://crispr.hzau.edu.cn). A fourth gRNA sequence from Feng et al., 2016 was used as a positive control since it was demonstrated to yield albino and yellow plants (Table 2). Each gRNA sequence was cloned into a T7-gRNA vector (Table 6), synthesized *in vitro,* and tested for ability to direct DNA cleavage by purified Cas9 following protocols from Liang et al., 2018. A 2.3kB PCR amplicon of the *Lemon White 1* gene was amplified from ILP1 DNA using primers Lw1-F1 and Lw1-R2 (Table 5). *Lw1* amplicons were digested using the purified

Cas9 and synthesized gRNAs to determine if the designed guide sequences were sufficient to cut the target DNA *in vitro* (Figure 3B). Amplicons incubated with Lw1-gRNA2 show a 476bp band, demonstrating cleavage between Lw1-gRNA2's target site and the reverse primer. Lw1-gRNA3 and Lw1-gRNA4's target sites are located nearby the forward and reverse primer sites, so cleavage of this amplicon did not result in a significant downshift that could be visualized on the gel. Mixing both Lw1-gRNA2 and Lw1-gRNA4 produced a similar pattern to Lw1-gRNA2. Finally, adding Lw1-gRNA2 and Lw1-gRNA3 reveals a 387bp band resulting from the further cleavage of the 476bp fragment generated by Lw1-gRNA2.

Mutations in the *Lemon White 1* gene in Illinois Low Protein and H99 inbred genotypes

The DNA sequences encoding each of the four Lw1-gRNAs were cloned into the pMOD_B2103 transformation vector, for expression as a single polycistronic transcript driven by the viral yellow cestrum promoter (Figure 4). This vector also expresses Csy4 as an N-terminal fusion to a Cas9 optimized for monocot expression, separated by the P2A self-cleaving peptide. Release of Csy4 enables processing of each individual functional gRNA. The assembled vector was used in biolistic transformation experiments with embryogenic calli from each of the H99, ILP1, and IHP1 genotypes. A number of transformed events were produced from each of the experiments with H99 and ILP1 calli at frequencies around 1% (Table 2), but none from IHP1 experiments of similar size.

Multiple albino plants were recovered during the regeneration stage of tissue culture, indicative of a biallelic mutation at the *Lemon White 1* locus (Figure 5A). PCR Amplification of the 5' region containing Lw1-gRNA 4 showed no obvious deletions and produced an expected band of 447bp (Figure 5B). This indicates that no large deletion occurred that spanned the entire gene. Sanger sequencing of these amplicons show a 6bp biallelic deletion in ILP Event 2 (ILP-E2) at the expected Cas9 cleavage site. Chromatograms of ILP Event 1 (ILP-E1) and H99 Event 1 (H99-E1) show decomposition at the Cas9 cleavage site when sequencing with the reverse primer (Figure 5E). This is indicative of heterozygous edits in these events and resequencing of the individual alleles is ongoing. PCR amplification of the 3' end of the gene using Lw1-F2 and Lw1-R2 primers revealed sizable deletions in ILP-E1 and (H99-E1) (Figure 5B). Since this region contained three guide RNAs, sizable deletions were to be expected due to simultaneous cleavage of the gene. DNA sequencing of PCR amplicons showed mutations induced by DSB repair and deletions between two guide RNAs (Figure C, D). The ILP-E1 produced an amplicon ~500 bp in length, a size consistent with the expected loss of 385 bp between Lw1-gRNA2 and Lw1-gRNA3. In the same ILP-E1 event, another amplicon at ~800 bp was produced, the result of small indels generated at Lw1-gRNA1 and Lw1-gRNA2, as well as a 29 bp deletion near the target site for Lw1-gRNA3. PCR of Lw1 in ILP-E2 produced a similar fragment as wildtype. Sequencing of this fragment revealed a 28 bp insertion in between Lw1-gRNA1 and Lw1-gRNA2, and small indels 3-4 bps upstream of their respective PAM sequences.

DISCUSSION

Illinois Long Term Selection Lines as transformation genotypes

Transformation and genome editing are powerful tools for advancing breeding and understanding unique germplasm such as the ILTSE. Biotechnology strategies enable the study of these lines, offering new opportunities for validating genes of interest that can improve complex agricultural traits such as Nitrogen Use Efficiency. Biolistic transformation yielded transgenic and edited ILP1 and H99 type I callus. In contrast, a majority of maize CRISPR experiments utilize type II callus genotypes, morphogenic regulators, or *Agrobacterium*-mediated transformation. The number of regenerated, edited ILP1 plants occurred at a ~1.5% transformation efficiency (Table 1), which is in the expected range when using biolistic methods. The absence of regenerated plants in the IHP1 background is noteworthy compared to the success of editing ILP1. IHP1 and ILP1 are both derived from an open pollinated variety called "Burr's White" but have diverged due to a result of selection for grain protein concentration. The lack of regenerants from IHP1 could be due to altered N metabolism or a differential response to antibiotics during the selection stage. The distinct response to auxins and callus initiation between ILP1 and IHP1 (Figure 2A, B) support the idea that media optimization might be required to regenerate transgenic IHP1 plants. Alternatively, the use of morphogenic regulators (Lowe et al., 2016) could aid in the generation of transgenic and edited IHP1 plants.

Editing pipelines for successful genome editing

Our goal was to develop a complete protocol for transforming and editing inbred lines

from the ILTSE. Prior to transformation, there were a number of steps that contributed to the pipeline. The first was gRNA validation *in vitro* using purified Cas9, synthesized gRNA molecules, and PCR amplicons of the target gene. Successful cleavage of PCR amplicons indicated that the designed gRNA was able to target the intended gene of interest and was advanced to plasmid cloning. Once the transformation vectors were complete, testing in maize B73 protoplasts can be performed to ensure successful *in vivo* gRNA cleavage and Csy4 processing. Multiplexing experiments facilitate the possibility of deleting a DNA sequence flanked by two gRNAs. PCR amplification of CRISPR transfected protoplasts revealed deletions easily seen on an agarose gel. Other methods of detecting CRISPR indel events include the T7 endonuclease assay, designing a mutation to eliminate a restriction enzyme site, or using tools such as TIDE (Brinkman et al., 2014). Screening for deletions greater than 40 basepairs by PCR saves time by eliminating the further steps needed to perform the mentioned methods. Similar deletions can be seen when screening populations of putative edited plants. One drawback to this method is that deletions can occur at two gRNA sites without loss of sequence between them. Another way to detect edits in this situation is to use the RNP made for performing *in vitro* cleavage assays as a tool to screen PCR amplicons from regenerated plants (Liang et al., 2018). These methods are a robust protocol for future editing of ITLSE lines and other inbreds to recover plants with mutations at multiple loci at an efficient frequency.

Csy4 as a strategy for multiplexing in maize

CRISPR RNAs containing 20 nucleotide target sequence and gRNA scaffold were placed

under control of the strong, constitutive *Cestrum* Yellow Leaf Curling Virus (CmYLCV) promoter (Stavolone et al., 2003). Most CRISPR literature in maize has utilized pol II promoters such as U3 and U6 from different monocots, which also require a 5' "G" or "A" nucleotide for efficient expression (Zhang et al., 2017). Using the Csy4 system, each gRNA is spaced with a 15bp Csy4 recognition site, removing the need for a 5' "G" nucleotide in each gRNA and allowing expression of multiple gRNAs using a single promoter. This gives more flexibility when designing guide RNAs for editing in maize. Given that each guide RNA site produced edits at their expected cut site distal to the NGG PAM sequence, this system is sufficient in processing each gRNA *in vivo* without requiring specific nucleotides for strong expression. It is also indicative that Csy4 can efficiently process four guide RNAs that are functional for mutagenesis. Compared to other studies that targeted the *Lemon White 1* locus, a variety of phenotypes were seen. In Feng et al., 2016 and 2018, using the single guide Lw1-gRNA1 yielded albino, yellow, chimeric, and wildtype *Lw1* phenotypes. Our study with the same Lw1-gRNA1 showed only albino or wildtype phenotypes, which is likely due to the fact that four gRNAs were used to knock out this gene, biasing our results to biallelic mutations. Monoallelic mutations are still possible given that Lw1-gRNA 4 produced heterozygous edits in ILP E1 and H99 E1.

Previous experiments were conducted with ILP1 and IHP1 utilizing self-cleaving ribozymes adapted from Tang et al., 2016; however, no edits were detected as determined solely by the appearance of the albino phenotype. (data not shown). Fertile transgenic plants positive for both the NPTII and Cas9 transgenes were recovered in ILP1,

demonstrating the ability for ILP1 to be transformed. The lack of mutated alleles from the initial experiments could be explained by the significantly lower efficiency of ribozymes compared to either the tRNA or Csy4 arrays (Cermak et al., 2017).

The observed mutations at multiple gRNA target sites demonstrates successful processing of individual gRNA *in vivo* with Csy4. This strategy is ideal for using multiple gRNAs targeting one gene or for simultaneous editing of multiple genes using a single promoter, thus allowing for smaller plasmid size and improving transformation efficiency. Study of the ILTSE has resulted in many candidate genes for functional genomics (Goldman et al., 1993, Moose et al., 2004), which can all be candidates for site-specific genome editing in the future. Editing one gene at a time may not yield a significant phenotype for traits such as N use and protein accumulation, making multiplexing an ideal strategy for mining the ILTSE for genes useful to agriculture. With the success of *Lemon White 1* proof of concept, the same protocol can be adapted for genome editing ILP1.

FIGURES AND TABLES

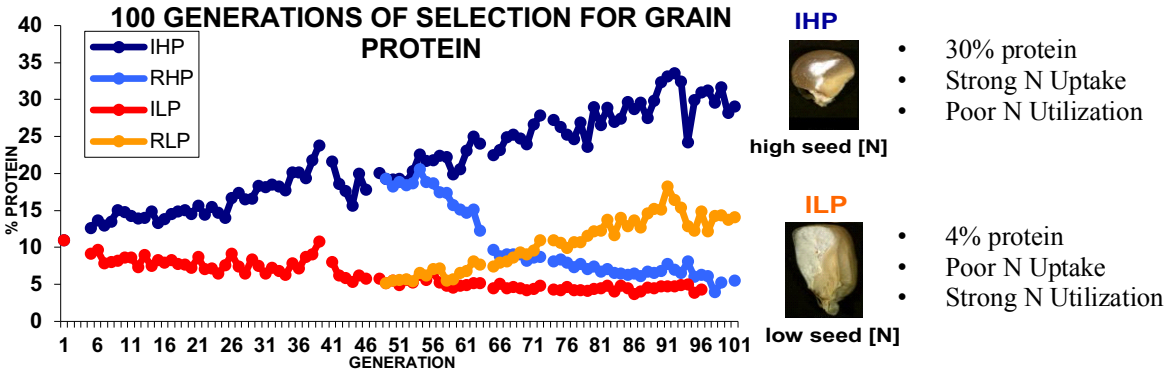

Figure 1. The Illinois Long Term Selection Experiment selection response for protein strains (adapted from Moose et al., 2004)
Protein concentration through generations of long term selection. Protein concentrations at generation one were within normal ranges of grain protein in kernels (8-12%) and have responded to directional selection of grain protein. Selection for high protein has resulted in the Illinois High Protein (IHP) population which is a high nitrogen (N) seed containing 30% grain protein. Conversely, low protein selection has resulted in the Illinois Low Protein (ILP) population at near 4% grain protein. Both populations contrast in their ability to uptake and utilize N. The reverse selection lines, Reverse High Protein (RHP) and Reverse Low Protein (RLP), were created by changing the direction of selection to study remaining genetic variability. Inbred lines used for tissue culture are derived from self pollinating IHP or ILP at cycle 90.

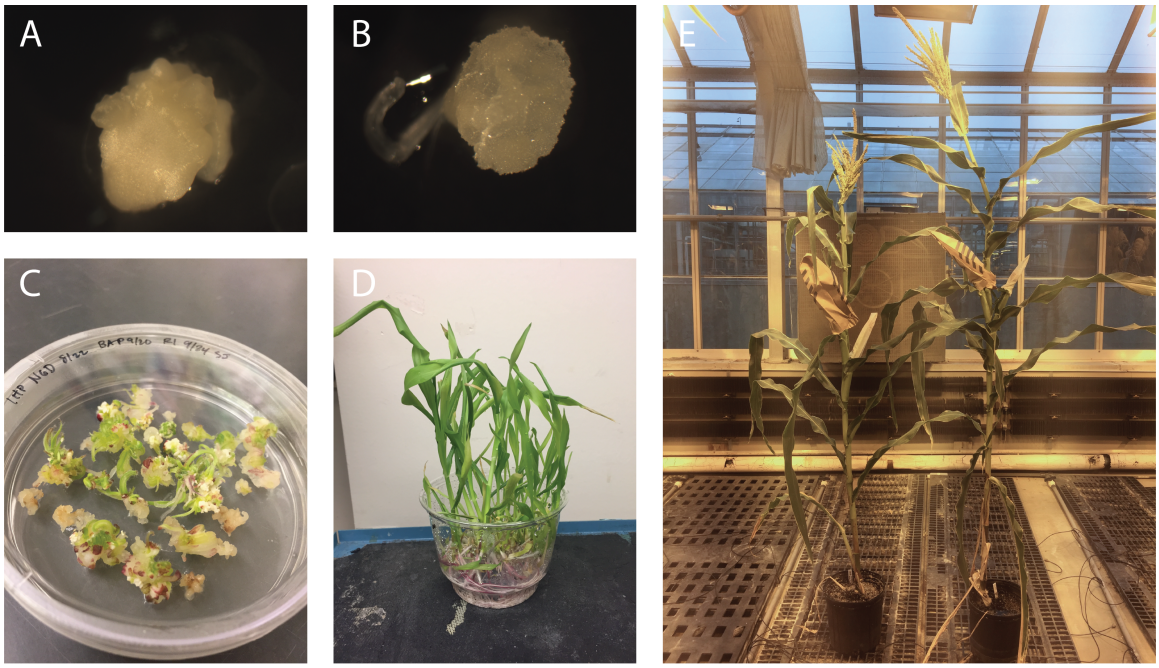

Figure 2. Tissue Culture of Illinois Protein Strains
(A) Illinois Low Protein 1 (ILP1) callus 10 days after embryo induction on N6 Dicamba media **(B)** Same as (A) for Illinois High Protein 1 (IHP1) **(C)** Regeneration of plantlets in light conditions on R1 MS based media following a 5 mg/L 6-BAP pulse for four days. **(D)** Continued regeneration of plantlets on R2 MS based media **(E)** Tissue cultured derived plants grown to flowering in greenhouse conditions. IHP1 (left) and ILP1 (right) both produced viable tassels and silks for further seed propagation by self pollination.

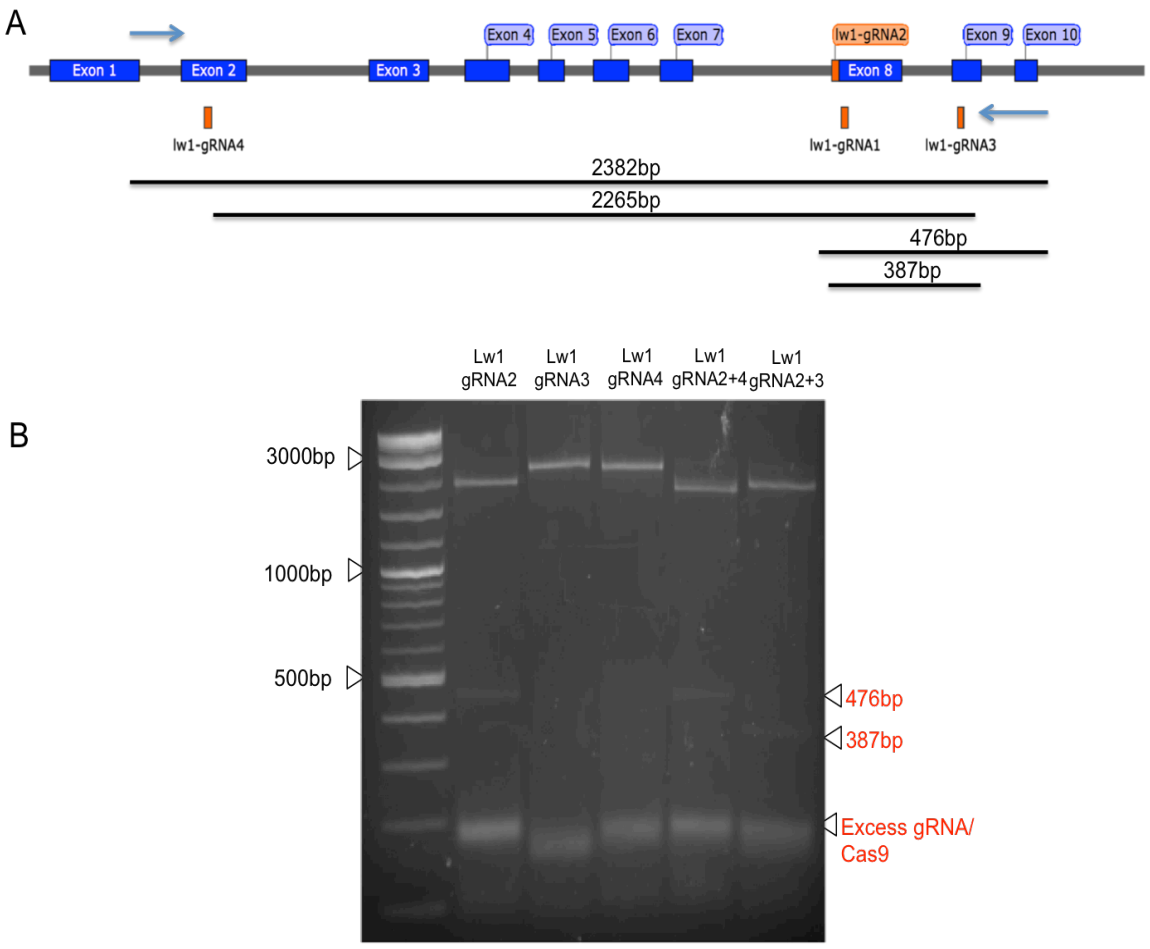

Figure 3. Testing of designed *Lemon White 1* gRNAs *in vitro*
(A) Lemon White 1 *Zm00001d03389* gene model depicting exons (in blue) and introns (in grey). Guide RNAs designed with CRISPR-P v2 or CHOPCHOP plotted on the model in orange. Primer locations made for PCR amplification of the gene are designated by arrows flanking the gene. Digested amplicon sizes and their expected sizes with the tested guide RNAs are shown below the gene model. **(B)** PCR amplicons of *Lemon White 1* digested with Cas9 and guide RNA. Lanes are labeled with each guide or guides added to each reaction.

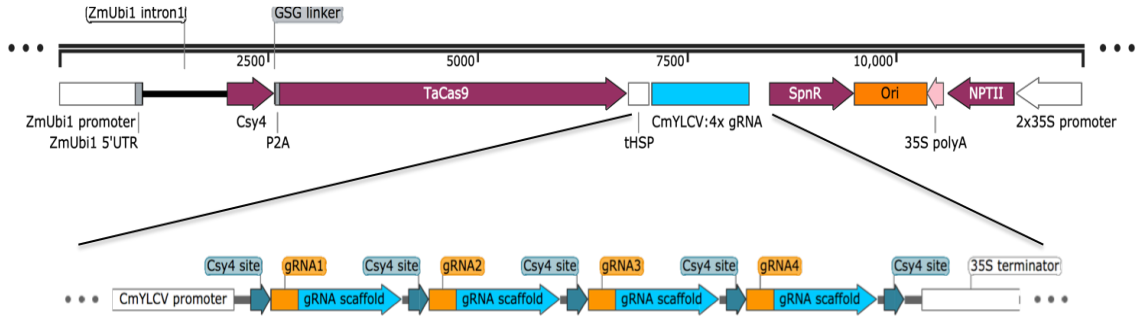

Figure 4. Transformation vector used for biolistic transformation
Plasmid map of transformation vectors built using protocols from Cermak et al., 2017 containing functional cassettes for plant tissue culture and editing. Cas9 and Csy4 linked with a P2A self cleaving peptide are expressed under the maize ubiquitin 1 promoter. Four guide RNAs spaced with Csy4 sequences were under the viral cestrum promoter. NPTII selectable marker was expressed using a 2x35S promoter.

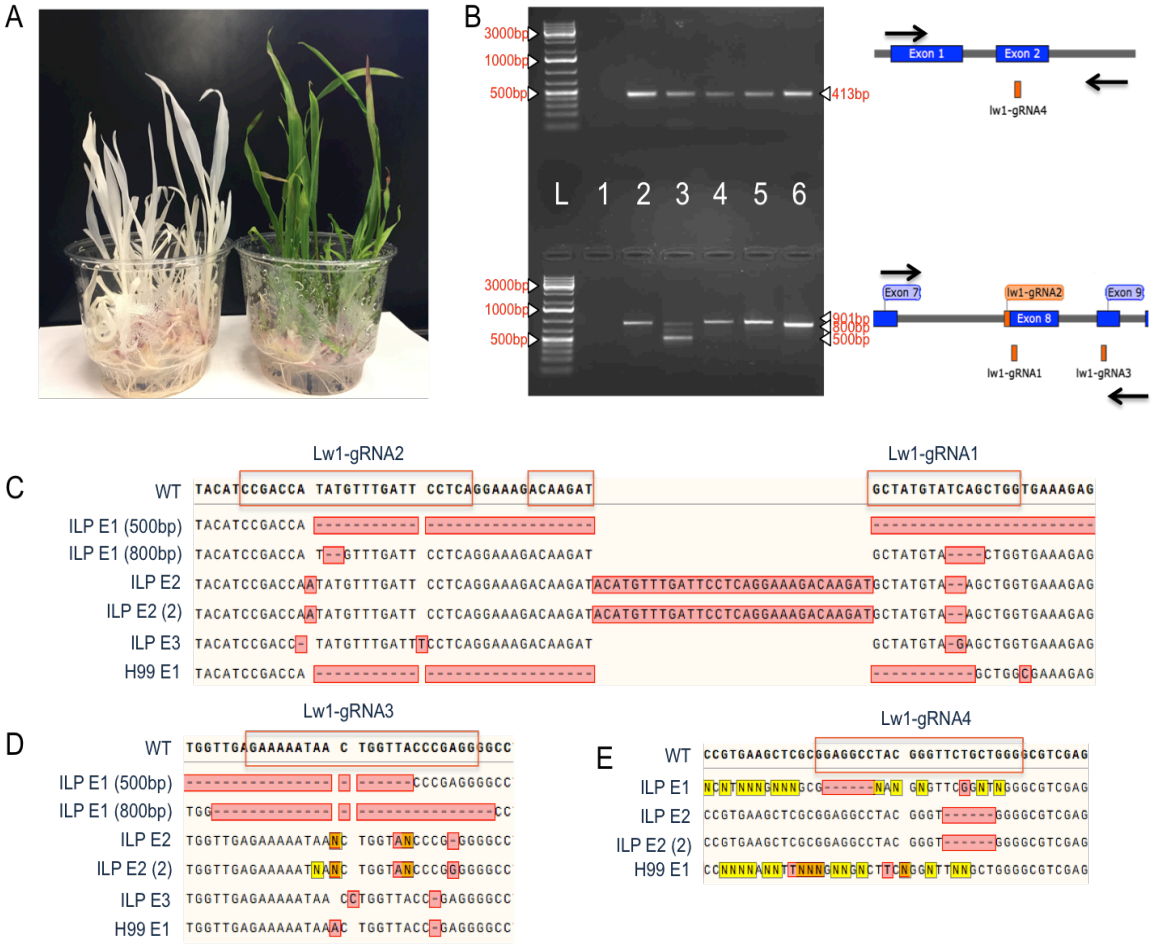

Figure 5. Characterization of edited *Lemon white 1* mutants
(A) Regenerated albino plant phenotype compared to wild-type **(B)** PCR amplification of plants with albino phenotype. Lanes L-2 are the NEB 1kb+ ladder, water negative control, and ILP1 wild-type respectively. Lanes 3-6 correspond to individual albino events sampled in the regeneration stage of tissue culture; Lane 3: ILP E1, Lane 4: ILP E2, Lane 5: ILP E2 (2), and Lane 6: H99 E1. The top half of the gel shows amplification of the 5' end of the gene targeted with lw1-gRNA4 designated by the adjacent gene model. Arrows on the gene model designate primers used for PCR. The bottom half of the gel depicts the 3' end of the gene with regions that lw1-gRNA1, 2, and 3 target. **(C/D/E)** Sanger sequencing of PCR products from the gel in (B). Corresponding guide RNA targets including PAM sequence are outlined in orange boxes.

Gene Name	Target Sequence (5' -> 3')	Abbreviation
Lemon White 1 (Zm00001d033896)	ACAAGATGCTATGTATCAGCTGG	lw1-gRNA1
	CCGACCATATGTTTGATTCCTCA	lw1-gRNA2
	GAAAAATAACTGGTTACCCGAGG	lw1-gRNA3
	GGAGGCCTACGGGTTCTGCTGGG	lw1-gRNA4

Table 1. Guide RNA design for *Lemon White 1*
Guide RNAs gene targets with corresponding gRNA sequences and abbreviations for gene models. The 20 nucleotide target sequence is labeled in black and the PAM sequence in red.

Experiment	# embryos	# NPTII resistant calli	#events regenerated	#plants regenerated	#plants with edits	Efficiency
H99 - Lw1	282	24	10	15	0	0%
ILP1- Lw1	226	78	16	30	0	0%
IHP1 - Lw1	244	0	0	0	0	0%
H99 - Lw1 Rep 2	252	32	14	15	1	0.4%
ILP1 - Lw1 Rep 2	202	280	15	22	3	1.49%
IHP1 - Lw1 Rep 2	204	0	0	0	0	0%

Table 2. Transformation Data in *Lemon White 1* experiments
Transformation frequencies of editing experiments in the ILTSE. *Lw1* experiments utilized a vector from Figure 4 with four guides for *Lemon White 1*. The number of embryos used designates how many were used during bombardment. NPTII resistant calli were determined by their growth on N6 media containing paramomycin (300mg/L). Events generated was determined by NPTII calli that turned into plantlets with shoots and roots during regeneration. The plants regenerated was determined by counting the number of plants survived regeneration and were transplanted into the greenhouse. Edited plants were determined by PCR. Efficiency was calculated as the number of plants with edits divided by the number of embryos induced for the experiment.

CHAPTER 2: CRISPR-CAS9 MEDIATED MULTIPLEX EDITING OF *L-ASPARAGINASE*

INTRODUCTION

Nitrogen Use and Metabolism in Maize

Cereal crops account for a majority of humankind's caloric intake and dominate global agriculture. Knowledge of nitrogen (N) uptake, utilization, and metabolism in cereal crops is crucial to understanding how we can satisfy the nutritional needs of a growing world population while minimizing environmental impacts from nitrogen fertilizers. Increases in population require more productivity from agriculture, and similarly more N, to keep up with the demand. Some estimates suggest that 45-73% more N will be required to reduce yield gaps (Beatty and Good, 2018). However, this increase in added N will come as a detriment to the environment (Han et al., 2015). Therefore, creating crops that can perform better under low N or are able to uptake and utilize N at a greater efficiency are essential for future agricultural production. The ability to produce high yield at a lower N rate is termed "Genetic N Efficiency" and will become a more important breeding goal heading into the future (Han et al., 2015). An important factor of N metabolism is accumulation in economic portions of the plant. This is termed Nitrogen Use Efficiency (NUE), and is described as the ratio of grain yield to the amount of N applied to the field (Moose and Below, 2009). NUE is a highly complex trait controlled by the genetic ability of a plant to assimilate inorganic forms of N from the soil and to shuttle organic forms to sink tissues. Environmental factors such as soil fertility and the soil microbiome also influence this trait. It is also related to other measures of N metabolism like N uptake, assimilation, and remobilization.

Inorganic N enters plants through the roots as nitrate or ammonia. Nitrates must first be reduced into ammonia before being incorporated into organic N forms such as amino acids. In maize, most N assimilation occurs in the leaves, where carbon acceptors for N from photosynthesis are generated (Moose and Below, 2004). The well-studied GS/GOGAT cycle has been reported as the primary N assimilation pathway in plants and yields aspartate and glutamine (Cren and Hirel, 1999). Synthesis of these amino acids occurs in the plastid during light conditions, while other downstream amino acids are synthesized in the dark. This includes asparagine, which is created from aspartate and glutamine by the enzyme asparagine synthetase (AS, Figure 6). Overexpression of AS in *Arabidopsis* elevated levels of soluble seed protein and improved plant tolerance to low N-containing media (Lam et al., 2003). Asparagine is also implicated as a storage amino acid and transporter of N due to its limited use as a metabolic substrate and high N to C ratio. It also functions as a signal for the vegetative N status to developing seeds in maize (Seebauer et al., 2004). N is remobilized from source to sink tissues as amino acids including aspartate, asparagine, glutamate, and glutamine. In these tissues, amino acids accumulate in storage proteins such as the zeins and globulins, which contribute to overall NUE.

Variation in the Asparagine Cycling Pathway in the Illinois Long Term Selection Experiment

The Illinois Long Term Selection Experiment (ILTSE) is a unique resource for studying genomic responses to selection and changes in whole plant N metabolism. The Illinois Protein Strains contrast drastically in their ability to utilize N. Mainly, the Illinois High

Protein (IHP) lines have an increased ability to assimilate N from the environment by having a larger root system. However, since this line hyperaccumulates N as storage proteins in the kernels, it cannot store as much starch as compared to most strains of corn. Therefore, overall N utilization is decreased due to smaller seed size. In contrast, Illinois Low Protein (ILP) lines develop significantly smaller root systems, thereby limited N uptake from the soil; however, these lines can better utilize N compared to IHP by accumulating starch in the seeds and having higher grain yield.

Likewise, genes related to N use in the asparagine cycling pathway have diverged due to long-term selection and make excellent candidates for functional genomics studies (Lucas 2014, PhD dissertation). The two main enzymes involved in asparagine cycling are Asparagine Synthetase (*AS*) and L-Asparaginase (*ASNase*). AS yields asparagine and glutamate by the amidation of aspartate with a N donor from glutamine (Figure 6). Conversely, *ASNase* deaminates asparagine and results in aspartate and a free ammonium. *ASNase* is annotated on the B73v4 reference genome (Jiao et al., 2017) as "Zm00001d002052" and information about this gene is obtainable on MaizeGDB (www.maizegdb.org). B73 RNA-seq data available from this database shows high expression in leaf tissues and early seed development (Walley et al., 2016).

There are two unique alleles of *ASNase* found in the ILTSE, one present at high frequencies in ILP, and the other nearly fixed in IHP (Figure 7) The IHP allele of *ASNase* is expressed at lower levels compared to the ILP allele (Figure 6). Interestingly, Lucas (2014) showed that when the direction of phenotypic selection was reversed in both IHP

and ILP to create Illinois Reverse High Protein (IRHP) and Illinois Reverse Low Protein (IRLP), the frequencies of *ASNase* alleles also shifted. In IRLP, cycle 69 shows this population having a high frequency of the ILP-like *ASNase* allele, which is consistent with IRLP being derived from ILP. However, by cycle 100, the IRLP population became fixed for the IHP-like *ASNase* allele. The converse is also true for IRHP, where the allele frequency shifts from being IHP-like to ILP-like during that period of selection (Lucas 2014, PhD dissertation).

Gene Editing of ILP *ASNase* to mimic IHP expression

Efforts to characterize the highly expressed *L-Asparaginase* (*ASNase*) allele from ILP1 have been made by creating near isogenic lines (NILs) with IHP1 (Arp, 2017). An ILP1 NIL was made by introgressing a region containing the IHP1 *ASNase* and 93 total genes. Conversely, an IHP1 NIL with ILP1 *ASNase* contained 505 genes within the introgression. NIR data comparing the ILP1: *ASNase-IHP1* NIL to ILP1 showed a slight increase in protein of 0.5% in 2015. In *Arabidopsis*, mutants in *ASNase* created by T-DNA insertions had elevated levels of free asparagine in their seed and did not suffer a small yield penalty under low N conditions (Ivanov et al., 2012). Coincidentally, IHP has significantly higher levels of free asparagine compared to ILP (Lohaus et al., 1998), which is consistent with the weak expression of *ASNase* as shown by both qPCR (Figure 6) and RNAseq data (Figure 7) generated in the Moose Laboratory. Thus, IHP1 can be considered to harbor a mutant *ASNase* with reduced function. Although the exact mutations causing lower expression of the IHP1 *ASNase* are unknown,sequence analysis

(Figure 7) suggests the deletion approximately 1000-bp upstream of the transcription start site could be responsible. .

Whereas NILs are a valid way to study gene function, newer technologies in gene editing are more precise and require less time to get results. This is dependent on reliable transformation, which we have now established for the ILP1 background. Since both N use and protein concentration are quantitative traits, differences in other genes in the IHP and ILP NILs could contribute to the observed changes in grain protein concentration. Utilizing the CRISPR-Cas9-Csy4 pipeline developed for the *Lemon White 1* work described in Chapter 1, *ASNase* can be targeted to make a knockout of ILP-*ASNase*. The objective of this study is to create knockout, IHP-like *ASNase* alleles in the ILP background with CRISPR Cas9 and Csy4 multiplexing.

MATERIALS AND METHODS

All methods described in Chapter 2 are consistent with the protocols followed in chapter one. All guide RNAs, primers, and oligonucleotides were created and designed for *L-Asparaginase*.

RESULTS

Guide RNA design and validation *in vitro*

Four guide RNAs for *L-Asparaginase (Zm00001d020052)* were identified from version 4 of the B73 reference genome (Jiao et al., 2017) using CHOPCHOP (http://chopchop.cbu.uib.no) and CRISPR-P 2.0 (http://crispr.hzau.edu.cn) (Table 3).

Each gRNA sequence was cloned into a T7-gRNA vector and synthesized *in vitro* using T7 RNA polymerase (Table 8). No clones containing *ASNase*-gRNA5 were found, so it was not tested *in vitro*. A 2.9kB PCR amplicon of the *ASNase* gene was amplified from ILP1 DNA using primers 5'-CCGCACACATAGAGACAGAG-3' and 5'-TCATCGACAACACCTGTGAC-3' (Table 7). *ASNase* amplicons were digested using the purified Cas9 and synthesized gRNAs to determine if the designed guide sequences were sufficient to cut the target DNA *in vitro* (Figure 8). Amplicons treated with *ASNase*-gRNA4 or *ASNase*-gRNA2 do not show an observable downshift compared to undigestedamplicons because their cleavage sites are close to theends of the amplicons. Digestion with *ASNase*-gRNA3 shows three distinct bands: undigested DNA of 2.9kb, a 2311bp band, and a 627bp band. When mixing all three guides together, a pattern similar to *ASNase*-gRNA3 is generated ; however, an extra 428bp band is also observed. This is a result of the 627bp band also being cleaved by *ASNase*-gRNA2. These results demonstrate the designed gRNA sequences are able to target *ASNase* amplicons and may cause a loss-of-function mutation *in vivo*.

Recovery of ILP1 lines with mutations in *ASNase*

Regenerated plants from this experiment were screened by PCR for mutations in the *ASNase* coding sequence. Sanger sequencing of two events, ILP: ASN 1, and ILP: ASN 2 revealed deletions at the expected gRNA cut sites. Both ILP:*ASNase* 1 and ILP:*ASNase* 2 showed similar biallelic mutation patterns, indicating that they are two plants derived from the same callus (Figure 9). In these lines, a 217 bp deletion occurred between *ASNase*-gRNA5 and *ASNase*-gRNA3 cleavage sites. Downstream of this region, a 3bp

deletion was observed at the cut site of *ASNase*-gRNA2. Consistent with the *Lw1* experiment in chapter 1, Csy4 was able to process gRNAs *in vivo* in ILP1 and multiple gRNAs with Cas9 produced mutations that were detectable by PCR.

DISCUSSION

Recovery of T1 seed

Mutant plants generated in this experiment produced limited amounts of T1 seed for further study. Most plants in the ILP1 background did not create ears. When crossed to wildype plants, no seed was recovered. There are three likely explanation of why flowering was an issue during the growth of these plants: Csy4 expression in the cell, gene knockout of *ASNase*, or environmental stress during flowering. No fertile Csy4 plants have been identified so far. However, one cross, wildtype H99 with ILP: *ASNase* ASN 2, produced six seeds that will be planted at the UIUC Research and Education Center this year. This hybrid should contain a mutated, loss of function *ASNase* allele as well as either the wildtype H99 allele or a potential new allele due to continued activity of Cas9 and the 4 gRNAs. If unsuccessful, it is possible that Csy4 could be affecting the plant by cleaving mRNAs containing Csy4 sites from native maize genes. This has been a hypothesis in the gene editing community that Csy4 cannot work in all plant species because Csy4 sites are found in essential genes, making the presence of this enzyme lethal. This has support as no regenerated plants containing edits with Csy4 constructs have been recovered from plants like Setaria and Sorghum (Colby Starker, personal communication). An alternative to using Csy4 would be to use the tRNA processing spacers between guide RNAs that utilize the host tRNA system to release the guide

RNAs. This has been demonstrated to yield fertile maize with multiple guide RNAs for editing (Qi et al., 2016).

Another reason contributing to low seed production is that an *ASNase* knockout could be lethal for pollen development and growth. Since *ASNase* genes are the primary mechanism for conversion of asparagine into asparatate, a loss of function mutant could interfere with N metabolism. Recent work in *Arabidopsis* demonstrated that knocking out the two predominant *ASNase* genes by T-DNA insertions did not impact seed production (Ivanov et al., 2012). Results from this model system shows that plants which lack *ASNase*s can develop normally, which suggests that crops like maize without *ASNase*s should be able to survive. Since IHP1 has reduced expression of this gene and propagates normally, it suggests that the drastically lower expression of *ASNase*s should not hinder development.

One additional explanation for the lack of seed production could have been due to heat stress in the greenhouse. During flowering, it was found that our greenhouse at the Plant Care Facility had abnormally high temperatures. At this time we had placed a number of plants in the greenhouse and thus turned on all of our growth lights. The control ILP1 plants grown alongside the transgenics did not produce seed, which provides support for this hypothesis.

Future Work

The data and results presented in this experiment were unable to produce a significant

amount of material for continued study. The recovery of novel *ASNase* alleles with this system provides valuable information for continued efforts to characterize *ASNase* function in the ILTSE. The same experiments could be repeated using plasmid pCas9-*ASNase*-NPTII to create knockouts of *ASNase* by targeting the coding region. However, new genomic data of both IHP1 and ILP1 could be useful for creating a more precise editing scheme for *ASNase*. Preliminary data suggests that the reduced expression of IHP-*ASNase* is caused by a series of mutations in the promoter region, which have been previously studied before in the Moose Lab by Farag Ibraheem and Yuhe Liu. Whole genome re-sequencing data is consistent with PCR based assays to distinguish between the IHP and ILP *ASNase* alleles, but the entire sequence is not known. Using primers that amplify the promoter region and primers designed for genotyping edits in *ASNase*, the 5' regulatory regions of the gene can be determined with Sanger sequencing. Knowledge of these regulatory regions can be used to design guide RNAs that can recreate the large deletion and indels within the IHP promoter region. This will mimic IHP-*ASNase* to a greater degree than just knocking out the coding sequence and could yield N phenotypes similar to the IHP lines. The resulting mutations from editing this region could also lead to other novel promoter sequences not seen in the ILTSE, such as the "promoter bashing" strategy that has created quantitative variation in tomato (Rodriguez-Leal et al., 2017).

The re-sequencing data also shows variation in the 3' UTR of *ASNase* between IHP1 and ILP1. ILP1 contains a deletion which could affect expression levels in this background. If a protocol for IHP1 transformation were established, then this region could be edited in that background to characterize this regions influence on gene function and expression.

Regenerated plants with altered *ASNase* can be phenotyped for amino acid content in leaves, developing earshoots, and seeds. Materials made from these experiments can also be analyzed for seed components like protein and oil by near infrared spectroscopy.

FIGURES AND TABLES

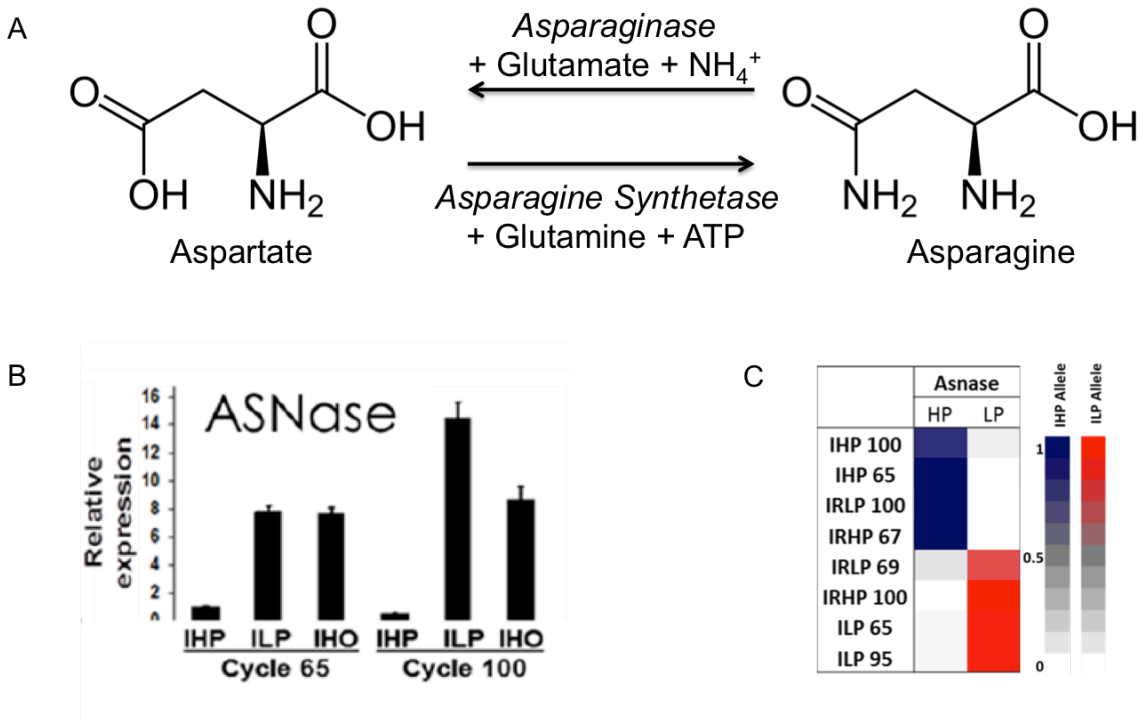

Figure 6. Asparagine Cycling and Expression in the Illinois Long Term Selection Experiment

(A) Amino acid metabolism focusing on Aspartate and Asparagine. Aspartate is converted to asparagine by *Asparagine Synthetase* with glutamine as a Nitrogen donor and ATP. Asparaginase converts asparagine back into aspartate by deamination with Glutamate. **(B)** Relative expression of *Asparaginase* in ILTSE lines at different generations determined by qPCR. Illinois High Protein (IHP) shows weak expression compared to Illinois Low Protein (ILP) and Illinois High Oil (IHO) and continues to go down through generations. Conversely, ILP has higher expression of *Asparaginase* compared to IHP and continues to increase when selecting for low grain protein. IHO *Asparaginase* is unchanged through generations. **(C)** Allele frequency of *Asparaginase* through the Illinois Protein Strains and Reverse Protein Strains. Blue designates IHP-like *Asparaginase* and red designates ILP-like *Asparaginase.* Data from Christine Lucas' PhD dissertation.

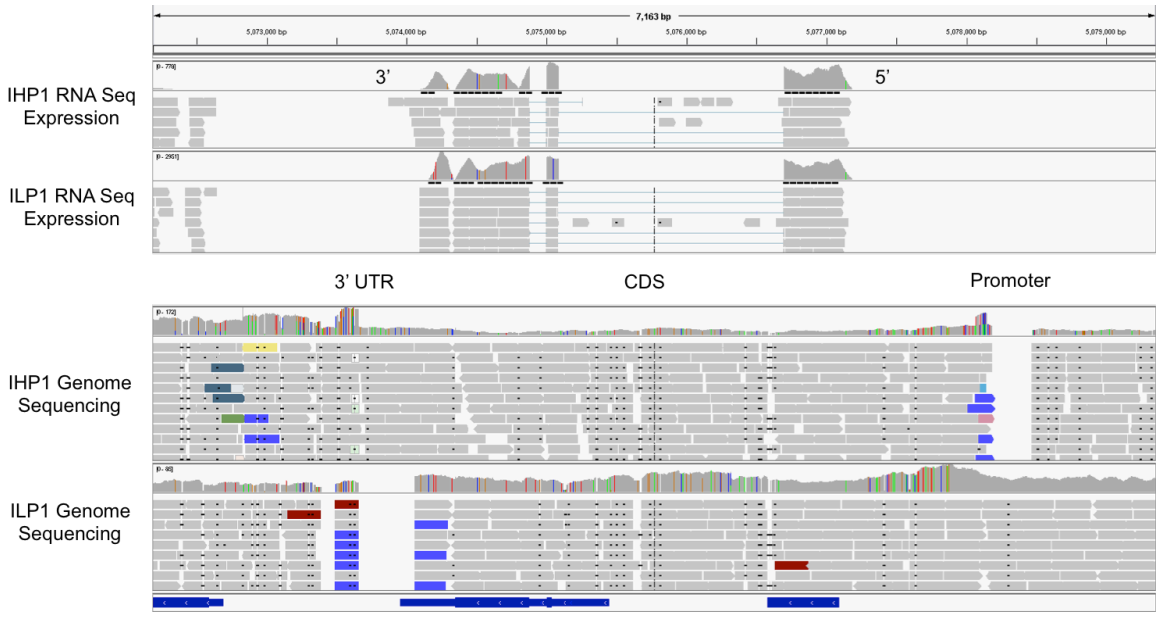

Figure 7. RNASeq and Genome Data of *Asparaginase* in Illinois High Protein 1 (IHP1) and Illinois Low Protein 1 (ILP1)
Genome data of Illinois Protein Strains displayed in Interactive Genomics Viewer (IGV). Each part of the gene is labeled in the center. The top tracks show RNAseq leaf libraries of both ILP1 and IHP1. Read count for *Asparaginase* in IHP1 is shown as [0-778] while ILP1 is [0-2951]. Each of the bottom tracks correspond to Illumina 400bp sequencing of the IHP1 and ILP1 genome.

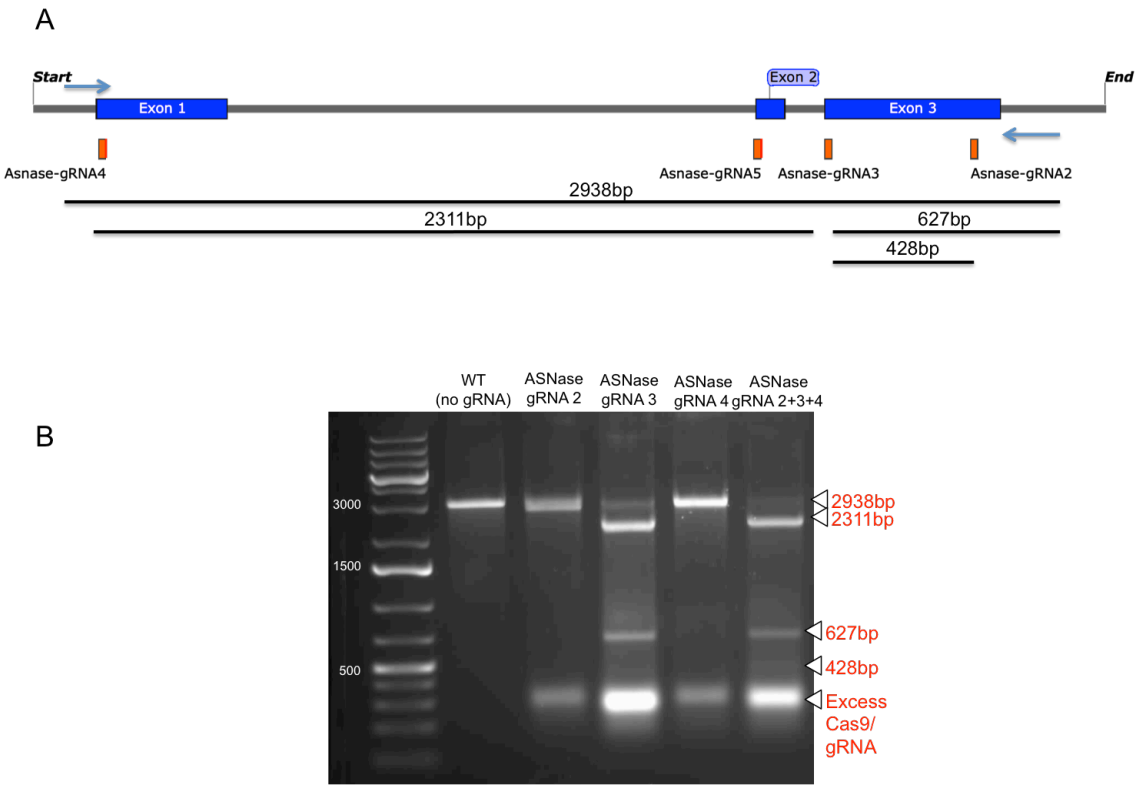

Figure 8. Testing of designed *L-Asparaginase* gRNAs *in vitro*
(A) Asparaginase *Zm00001d002052* gene model depicting exons (in blue) and introns (in grey). Guide RNAs designed with CRISPR-P v2 or CHOPCHOP plotted on the model in orange. Primer locations made for PCR amplification of the gene are designated by arrows flanking the gene. Digested amplicon sizes and their expected sizes with the tested guide RNAs are shown below the gene model. **(B)** PCR amplicons of *L-Asparaginase* digested with Cas9 and corresponding guide RNAs.

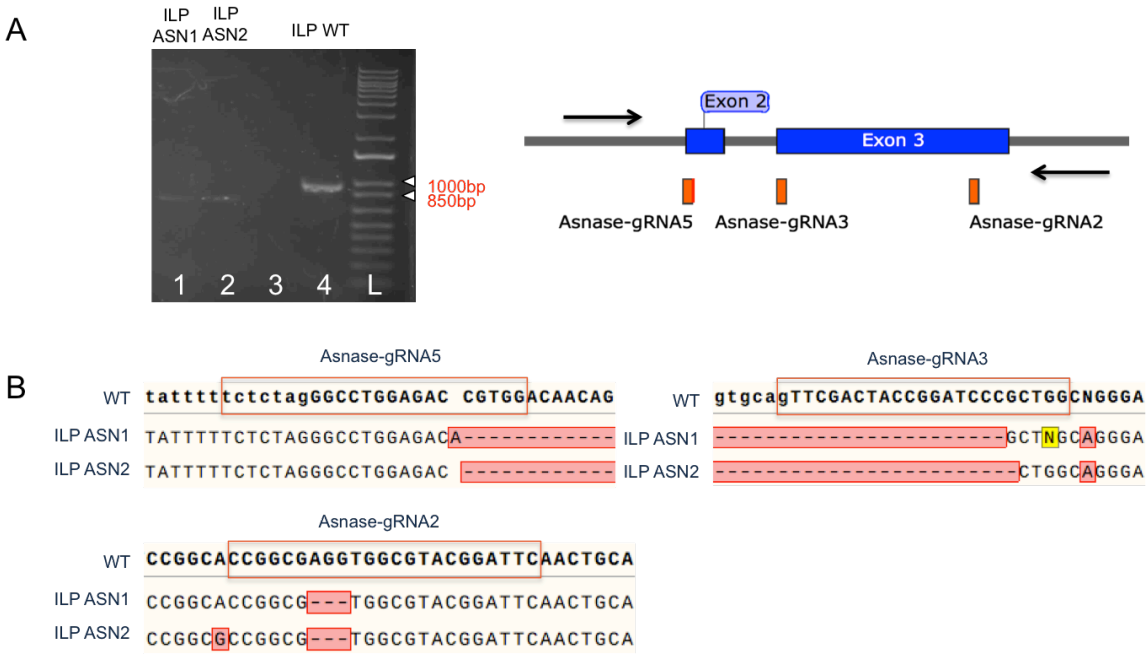

Figure 9. Characterization of edited *Asparaginase* mutants
(A) PCR amplification of the 3' end of *Asparaginase* containing exon 2 and 3. Lanes 1 and 2 are ILP ASN events 1 and 2, lane 3 and 3 are negative and wild-type control, and lane L is NEB 1kB+ DNA ladder. The gel shows amplification of the 3' end of the gene targeted with *ASNase*-gRNA 2, 3, and 5 designated by the adjacent gene model. Arrows on the gene model designate primers used for PCR. **(B)** Sanger sequencing of PCR products from the gel in (A). Corresponding guide RNA targets including PAM sequence are outlined in orange boxes.

Gene Name	Target Sequence (5' -> 3')	Abbreviation
Asparaginase (Zm00001d002052)	CCGGCGAGGTGGCGTACGGATTC	*ASNase*-gRNA2
	GTTCGACTACCGGATCCCGCTGG	*ASNase*-gRNA3
	TGGGCCATTGCCATCCACGGCGG	*ASNase*-gRNA4
	TCTCTAGGGCCTGGAGACCGTGG	*ASNase*-gRNA5

Table 3. Guide RNA design for *Asparaginase*
Guide RNAs gene targets with corresponding gRNA sequences and abbreviations for gene models. The 20 nucleotide target sequence is labeled in black and the PAM sequence in red.

Experiment	# embryos	# NPTII resistant calli	#Events regenerated	#Plants regenerated	#Plants with edits	Efficiency
H99 - ASN	289	Contaminated	0	0	0	0%
ILP1 - ASN	241	238	72	98	0	0%
H99 - ASN Rep 2	267	112	14	36	0	0%
ILP1 - ASN Rep 2	212	56	17	33	2	0.94%

Table 4. Transformation Data for *Asparaginase* experiments
Transformation frequencies of editing experiments in the ILTSE targeting *Asparaginase* (ASN). This table utilizes the same criteria as Table 2.

REFERENCES

Armstrong, C. L., Green, C. E., & Phillips, R. L. (1991). Development and availability of germplasm with high Type II culture formation response. *Maize genetics cooperation news letter (USA)*.

Arp, J. Discovery of novel regulators and genes in nitrogen utilization pathways in maize. 2017. University of Illinois Urbana-Champaign, PhD dissertation.

Beatty, P. H., & Good, A. G. (2018). Improving Nitrogen Use Efficient in Crop Plants Using Biotechnology Approaches. In *Engineering Nitrogen Utilization in Crop Plants* (pp. 15-35). Springer, Cham.

Brinkman, E. K., Chen, T., Amendola, M., & van Steensel, B. (2014). Easy quantitative assessment of genome editing by sequence trace decomposition. *Nucleic acids research*, *42*(22), e168-e168.

Doudna, J. A., & Charpentier, E. (2014). The new frontier of genome engineering with CRISPR-Cas9. *Science*, *346*(6213), 1258096.

Cermak, T., Curtin, S. J., Gil-Humanes, J., Čegan, R., Kono, T. J., Konečná, E., ... & Voytas, D. F. (2017). A multi-purpose toolkit to enable advanced genome engineering in plants. The Plant Cell, tpc-00922.

Cren, M., & Hirel, B. (1999). Glutamine synthetase in higher plants regulation of gene and protein expression from the organ to the cell. *Plant and Cell Physiology*, *40*(12), 1187-1193.

Duncan, D. R., Williams, M. E., Zehr, B. E., & Widholm, J. M. (1985). The production of callus capable of plant regeneration from immature embryos of numerous Zea mays genotypes. Planta, 165(3), 322-332.

Feng, C., Yuan, J., Wang, R., Liu, Y., Birchler, J. A., & Han, F. (2016). Efficient targeted genome modification in maize using CRISPR/Cas9 system. Journal of Genetics and Genomics, 43(1), 37-43.

Feng, C., Su, H., Bai, H., Wang, R., Liu, Y., Guo, X., ... & Han, F. (2018). High-efficiency genome editing using a dmc1 promoter-controlled CRISPR/Cas9 system in maize. Plant biotechnology journal.

Frame, B. R., Zhang, H., Cocciolone, S. M., Sidorenko, L. V., Dietrich, C. R., Pegg, S. E., ... & Wang, K. (2000). Production of transgenic maize from bombarded type II callus: effect of gold particle size and callus morphology on transformation efficiency. In Vitro Cellular & Developmental Biology-Plant, 36(1), 21-29.

Goldman, I. L., Rocheford, T. R., & Dudley, J. W. (1993). Quantitative trait loci influencing protein and starch concentration in the Illinois long term selection maize strains. *Theoretical and Applied Genetics*, *87*(1-2), 217-224.

Han, M., Okamoto, M., Beatty, P. H., Rothstein, S. J., & Good, A. G. (2015). The genetics of nitrogen use efficiency in crop plants. *Annual Review of Genetics*, *49*, 269-289.

Haurwitz, R. E., Jinek, M., Wiedenheft, B., Zhou, K., & Doudna, J. A. (2010). Sequence-and structure-specific RNA processing by a CRISPR endonuclease. *Science*, *329*(5997), 1355-1358.

Hopkins, C.G. 1899. Improvement in the chemical composition of the corn kernel. p. 205–240. Ill. Agric. Exp. Stn. Bull. 55.

Ishida, Y., Hiei, Y., & Komari, T. (2007). Agrobacterium-mediated transformation of maize. *Nature protocols*, *2*(7), 1614.

Ivanov, A., Kameka, A., Pajak, A., Bruneau, L., Beyaert, R., Hernández-Sebastià, C., & Marsolais, F. (2012). Arabidopsis mutants lacking asparaginases develop normally but exhibit enhanced root inhibition by exogenous asparagine. *Amino Acids*, *42*(6), 2307-2318.

Jiao, Y., Peluso, P., Shi, J., Liang, T., Stitzer, M. C., Wang, B., ... & Guill, K. (2017). Improved maize reference genome with single-molecule technologies. *Nature*, *546*(7659), 524.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., & Charpentier, E. (2012). A programmable dual-RNA–guided DNA endonuclease in adaptive bacterial immunity. *Science*, *337*(6096), 816-821.

Jones, T., Lowe, K., Hoerster, G., Anand, A., Wu, E., Wang, N., ... & Gordon-Kamm, W. (2019). Maize Transformation Using the Morphogenic Genes Baby Boom and Wuschel2. In *Transgenic Plants* (pp. 81-93). Humana Press, New York, NY.

Lam, H. M., Wong, P., Chan, H. K., Yam, K. M., Chen, L., Chow, C. M., & Coruzzi, G. M. (2003). Overexpression of the ASN1 gene enhances nitrogen status in seeds of Arabidopsis. *Plant physiology*, *132*(2), 926-935.

Liang, Z., Chen, K., Zhang, Y., Liu, J., Yin, K., Qiu, J. L., & Gao, C. (2018). Genome editing of bread wheat using biolistic delivery of CRISPR/Cas9 in vitro transcripts or ribonucleoproteins. *Nature protocols*, *13*(3), 413.

Liang, Z., Chen, K., Yan, Y., Zhang, Y., & Gao, C. (2018). Genotyping genome-edited mutations in plants using CRISPR ribonucleoprotein complexes. *Plant biotechnology journal*, *16*(12), 2053-2062.

Lohaus, G., Büker, M., Hußmann, M., Soave, C., & Heldt, H. W. (1998). Transport of amino acids with special emphasis on the synthesis and transport of asparagine in the Illinois Low Protein and Illinois High Protein strains of maize. *Planta*, *205*(2), 181-188.

Lowe, K., Wu, E., Wang, N., Hoerster, G., Hastings, C., Cho, M. J., ... & Wang, L. (2016). Morphogenic regulators Baby boom and Wuschel improve monocot transformation. *The Plant Cell*, *28*(9), 1998-2015.

Lucas, C. Genetic analysis of grain protein concentration in the illinois protein strain recombinant inbred population of maize. 2014. University of Illinois Urbana-Champaign, PhD dissertation.

Moose, S. P., Dudley, J. W., & Rocheford, T. R. (2004). Maize selection passes the century mark: a unique resource for 21st century genomics. *Trends in plant science, 9*(7), 358-364.

Moose, S., & Below, F. E. (2009). Biotechnology approaches to improving maize nitrogen use efficiency. In *Molecular genetic approaches to maize improvement* (pp. 65-77). Springer, Berlin, Heidelberg.

Porebski, S., Bailey, L. G., & Baum, B. R. (1997). Modification of a CTAB DNA extraction protocol for plants containing high polysaccharide and polyphenol components. *Plant molecular biology reporter*, *15*(1), 8-15.

Qi, W., Zhu, T., Tian, Z., Li, C., Zhang, W., & Song, R. (2016). High-efficiency CRISPR/Cas9 multiplex gene editing using the glycine tRNA-processing system-based strategy in maize. *BMC biotechnology*, *16*(1), 58.

Raji, J. A., Frame, B., Little, D., Santoso, T. J., & Wang, K. (2018). Agrobacterium-and Biolistic-Mediated Transformation of Maize B104 Inbred. In *Maize* (pp. 15-40). Humana Press, New York, NY.

Rodríguez-Leal, D., Lemmon, Z. H., Man, J., Bartlett, M. E., & Lippman, Z. B. (2017). Engineering quantitative trait variation for crop improvement by genome editing. *Cell*, *171*(2), 470-480.

Seebauer, J. R., Moose, S. P., Fabbri, B. J., Crossland, L. D., & Below, F. E. (2004). Amino acid metabolism in maize earshoots. Implications for assimilate preconditioning and nitrogen signaling. *Plant Physiology*, *136*(4), 4326-4334.

Shi, J., Gao, H., Wang, H., Lafitte, H. R., Archibald, R. L., Yang, M., ... & Habben, J. E. (2017). ARGOS 8 variants generated by CRISPR-Cas9 improve maize grain yield under field drought stress conditions. *Plant biotechnology journal*, *15*(2), 207-216.

Songstad, D. D., Armstrong, C. L., Petersen, W. L., Hairston, B., & Hinchee, M. A. W. (1996). Production of transgenic maize plants and progeny by bombardment of Hi-II immature embryos. *In vitro cellular & developmental biology-plant*, *32*(3), 179-183.

Stavolone, L., Kononova, M., Pauli, S., Ragozzino, A., de Haan, P., Milligan, S., ... & Hohn, T. (2003). Cestrum yellow leaf curling virus (CmYLCV) promoter: a new strong constitutive promoter for heterologous gene expression in a wide variety of crops. *Plant molecular biology*, *53*(5), 703-713.

Svitashev, S., Young, J. K., Schwartz, C., Gao, H., Falco, S. C., & Cigan, A. M. (2015). Targeted mutagenesis, precise gene editing, and site-specific gene insertion in maize using Cas9 and guide RNA. *Plant physiology*, *169*(2), 931-945.

Svitashev, S., Schwartz, C., Lenderts, B., Young, J. K., & Cigan, A. M. (2016). Genome editing in maize directed by CRISPR–Cas9 ribonucleoprotein complexes. *Nature communications*, *7*, 13274.

Tang, X., Zheng, X., Qi, Y., Zhang, D., Cheng, Y., Tang, A., ... & Zhang, Y. (2016). A single transcript CRISPR-Cas9 system for efficient genome editing in plants. *Molecular plant*, *9*(7), 1088-1091.

Uribelarrea, M., Moose, S. P., & Below, F. E. (2007). Divergent selection for grain protein affects nitrogen use in maize hybrids. Field Crops Research, 100(1), 82-90.

Walley, J. W., Sartor, R. C., Shen, Z., Schmitz, R. J., Wu, K. J., Urich, M. A., ... & Briggs, S. P. (2016). Integration of omic networks in a developmental atlas of maize. *Science*, *353*(6301), 814-818.

Xing, H. L., Dong, L., Wang, Z. P., Zhang, H. Y., Han, C. Y., Liu, B., ... & Chen, Q. J. (2014). A CRISPR/Cas9 toolkit for multiplex genome editing in plants. *BMC plant biology*, *14*(1), 327.

Zhang, T., Gao, Y., Wang, R., & Zhao, Y. (2017). Production of guide RNAs in vitro and in vivo for CRISPR using ribozymes and RNA polymerase II promoters. Bio-protocol, 7(4).

APPENDIX A: Supplemental Tables

Table 5. Primers to genotype *Lemon White 1* mutations

Name	Primer Sequence
lw1-F1	CACTTCATGGCCTTCAATAC
lw1-R2	ACCTTATCTGGAGTTGAGGCAC
lw1-F2	GTCATCAAGACGCTCAAGGAG
lw1-R5	TGCAGTTAAGGCACGAACAC

Table 6. Primers used to clone gRNAs into T7 plasmid for *In vitro* assay

Primer Name	Primer Sequence (5' -> 3')	gRNA
T7_lw1_g2_F	TAGGTGAGGAATCAAACATATGGT	lw1-gRNA2
T7_lw1_g2_R	AAACACCATATGTTTGATTCCTCA	lw1-gRNA2
T7_lw1_g3_F	TAGGGAAAAATAACTGGTTACCCG	lw1-gRNA3
T7_lw1_g3_R	AAACCGGGTAACCAGTTATTTTTC	lw1-gRNA3
T7_lw1_g4_F	TAGGGGAGGCCTACGGGTTCTGCT	lw1-gRNA4
T7_lw1_g4_R	AAACAGCAGAACCCGTAGGCCTCC	lw1-gRNA4
T7 Promoter F	TAATACGACTCACTATAGG	All
gRNA Scaffold R	AAAAGCACCGACTCGGTGCCACTT	All

Table 7. Primers used to genotype *ASNase* mutants

Name	Primer Sequence
ASN-F1	CCGCACACATAGAGACAGAG
ASN-R1	GTTTGTTAGGGGGTTTCGTT
ASN-F2	TTAGCTGGGGATTTTCAACAGT
ASN-R2	TCATCGACAACACCTGTGAC

Table 8. Primers used to clone *ASNase* gRNAs into T7 plasmid for *In vitro* assay

Name	Primer Sequence	gRNA
T7_ASN_g2_F	TAGGGAATCCGTACGCCACCTCGC	ASNase-gRNA2
T7_ASN_g2_R	AAACGCGAGGTGGCGTACGGATTC	ASNase-gRNA2
T7_ASN_g3_F	TAGGGTTCGACTACCGGATCCCGC	ASNase-gRNA3
T7_ASN_g3_R	AAACGCGGGATCCGGTAGTCGAAC	ASNase-gRNA3
T7_ASN_g4_F	TAGGTGGGCCATTGCCATCCACGG	ASNase-gRNA4
T7_ASN_g4_R	AAACCCGTGGATGGCAATGGCCCA	ASNase-gRNA4
T7 Promoter F	TAATACGACTCACTATAGG	All
gRNA Scaffold R	AAAAGCACCGACTCGGTGCCACTT	All

APPENDIX B: Protocols

T7-gRNA cloning strategy

Adapted from Jao et al. 10.1073/pnas.1308335110 Supplementary Information

Prework: Design oligonucleotides

5’ – TAGG – N20 gRNA sequence – 3’

5’ – AAAC – N20’ gRNA sequence – 3’

1. Anneal oligonucleotides containing target sequence and BsmBI overhang
    a. 5 min denature at 95C
    b. Cooling to 50C at -0.1C/second
    c. Pause at 50C for 10 min
    d. Cool to 4C at -1C/second
2. Golden Gate Reaction
    a. 1ul annealed oligos
    b. 400ng T7-gRNA cloning vector
    c. 0.5ul BsmBI
    d. 0.5ul SalI
    e. 0.5ul BglII
    f. 0.5ul T4 DNA Ligase
    g. 1ul 10x NEBuffer 3
    h. 1ul 10x T4 ligase buffer
    i. H20 to 10ul
3. Digest / Ligation Program
    a. 3x (20 min at 37C / 15 min at 16C)
    b. 10 min at 37C
    c. 15 min at 55C
4. Transform 2ul into DH5alpha

Library Construction for Amplicon Sequencing
Gene specific primers designed to amply a <500bp portion of a target gene of interest were ordered containing the Illumina overhang adapter sequences:

Forward overhang: 5' TCGTCGGCAGCGTCAGATGTGTATAAGAGACAG-[locus specific sequence]
Reverse overhang: 5'GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAG-[locus specific sequence].

These primers were used to amplify their respective target sequence from extracted callus DNA using New England Biolab's Phusion High-Fidelity PCR Master Mix with HF Buffer (Catalog # M0531S). The resulting PCR products were cleaned up using an Ampure XP Bead Kit from Beckman Coulter (Catalog # A63880). A second round of PCR was using a Nextera XT Index Kit from Illumina (Catalog # FC-131-1001) for barcoding each gene specific region. The final PCR products were cleaned up again using the Ampure XP beads and pooled for sequencing on an Illumina MiSeq 2x250nt PE NANO v2 lane in the Keck Center at UIUC.

Table 9. Oligonucleotides used to amplify regions near gRNA target sites for Illumina MiSeq Sequencing

Gene	Primer (5' -> 3')	Name	gRNAs contained	Amplicon Size
Asparaginase Zm00001d002052	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCCGCACACATAGAGACAGAG	ILL-ASN-F4	Asnase-gRNA4	413
	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGGAAGGCGAGGTACGAGTG	ILL-ASN-R4	Asnase-gRNA4	413
	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCAACAACACGGCAAGATCC	ILL-ASNg3-5-F3	Asnase-gRNA3 & 5	436
	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGAGCCCGTTCATGACCATC	ILL-ASNg3-5-R3	Asnase-gRNA3 & 5	436
	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGGGCTCATGAACAAGATGG	ILL-ASNg2-F2	Asnase-gRNA2	448
	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGTCATCGACAACACCTGTGAC	ILL-Asng2-R2	Asnase-gRNA2	448
Lemon White 1 Zm00001d033896	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGGTCATCAAGACGCTCAAGGAG	ILL-lw1-F2	lw1-gRNA 4	413
	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGTGCAGTTAAGGCACGAACAC	ILL-lw1-R5	lw1-gRNA 4	413
	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGTTTGACATATGCTGCATTGACTT	ILL-lw1-F3	lw1-gRNA 1,2,3	479
	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGACCTTATCTGGAGTTGAGGCAC	ILL-lw1-R2	lw1-gRNA 1,2,3	479
Extra Lw1 Primers	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGTGCTCCACTTTTGGTTTGAG	ILL-lw1-F6	lw1-gRNA 1 & 2	436
	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGCAATCCAATTTTCGGTTAAAATG	ILL-lw1-R6	lw1-gRNA 1 & 2	436
	TCGTCGGCAGCGTCAGATGTGTATAAGAGACAGCAAAGGATTGGACCAGGAAA	ILL-lw1-F7	lw1-gRNA 3	400
	GTCTCGTGGGCTCGGAGATGTGTATAAGAGACAGTGAAGAGCATCCTCAACAACC	ILL-lw1-R7	lw1-gRNA 3	400

Table 10. Tissue culture media for IHP1, ILP1, and H99.

N6 Media (N6)	***0.5L***	***1L***
N6 Salts + Vitamins	2 g	4 g
*2,4-D (10 mg/ml)**	100 uL	200 uL
*Dicamba**	1.65 mg	3.315mg
Myo-inositol	50 mg	100 mg
L-Proline	1.38 g	2.76 g
Casein Hydrosylate	50 mg	100 mg
Sucrose	15 g	30 g
pH = 5.8		
Phytagel	1.25 g	2.50 g
AgNO3 (after autoclave)	125 uL	250 uL
Paramomycin (after autoclave)	150mg	300mg
Pour in 10mm plates		

Regeneration 1 (R1)	**0.5L**	**1L**
4.4 MS salts + vitamin	2.2 g	4.4 g
Sucrose	30 g	60 g
pH= 5.8		
Phytagel	1.375 g	2.75 g
Paramomycin (for selection)	150mg	300mg
Pour in 25mm plates		

Regeneration 2 (R2)	**0.5L**	**1L**
4.4 MS salts + vit	2.2 g	4.4 g
Sucrose	15 g	30 g
pH = 5.8		
Phytagel	1.375 g	2.75 g
Paramomycin (for selection)	75mg	150mg
Pour in magenta boxes or solo cups		

N6 Osmotic (N6O)	***0.5L***	***1L***
N6 Salts + Vitamins	2 g	4 g
*2,4-D (10 mg/ml)**	100 uL	200 uL
*Dicamba**	1.65 mg	3.315mg
Myo-inositol	50 mg	100 mg
L-Proline	0.35g	0.69g
Casein Hydrosylate	50 mg	100 mg
Sucrose	15 g	30 g
Sorbitol	18.2g	36.4g
Mannitol	18.2g	36.4g
pH=5.8		
Phytagel	1.25g	2.50g

Regeneration 1 + BAP (R1-BAP)	**0.5L**	**1L**
4.4 MS salts + vitamin	2.2 g	4.4 g
Sucrose	30 g	60 g
6-Benzoaminopurine		
pH= 5.8		
Phytagel	1.375 g	2.75 g
Paramomycin (for selection)	150mg	300mg
Pour in 25mm plates		

Note: ILTSE lines use Dicamba as an auxin source and H99 uses 2,4-D as an auxin source

Tissue Culture of ILTSE Protein Strains and H99

1. Grow genotypes in field or greenhouse. IHP1 and ILP1 perform best in the greenhouse when planted in late summer and flower into the fall. H99 performs well year round in the greenhouse. Self or sib plants and write date of pollination on bag.

2. 10 days after pollination begin to check ears for embryo size by cutting the top third of a kernel with a razorblade. Remove both the endosperm and embryo using a microspoon and carefully separate the embryo from the endosperm. Measure the embryo size. When embryos are 1-2mm, remove the ear for culture initiation.
    a. PREWORK: Autoclave two 1L plastic containers, 1L water, and stainless steel forceps for culture initiation step. Create a 30% bleach solution using chlorox and Milli-Q water. Prepare culture initiation medium **N6D** (containing Dicamba for ILTSE lines) or **N6** (containing 2,4-D for H99) and **N6O** (for bombardment).

3. Break off the top few kernel rows of the ear to expose the cob. Place a scalpel handle in the cop to create a handle to hold the ears. Submerge each ear in a 1L plastic container full of 30% bleach for 15 minutes. After sterilizing, rinse each ear thoroughly with sterile water in another sterile 1L container to remove most of the bleach. Remove the top third of all kernels and use a sterile microspoon to remove the endosperm and embryo from the ear. Carefully separate the embryo from the endosperm and place the embryo with the flat side down on the media. Place plates in dark incubator at 25C, check for contamination regularly.
    a. Note: Maize embryos have a line on the flat side of the embryo. If necessary, check embryos with a dissecting microscope to ensure each embryo is in the correct orientation. Callus will not grow on upside down embryos.

4. Follow the bombardment protocol utilized for the Moose Lab when embryos or callus are at the correct stage. Bombardment of embryos saves time in tissue culture compared to callus.
    a. PREWORK: Prepare **N6** or **N6D** media containing 300mg/L Paramomycin for selection of transformants with the NPTII selectable marker

5. Allow embryos to rest for at least a day or up to 5 days after transformation on **N6** or **N6D** media without antibiotics. Do not leave tissues on **N6O** for more than a day.

6. Move transformed tissue to **N6** or **N6D** with antibiotics for selection and incubate in dark at 25C. Subculture each callus every two weeks for 6-12 weeks. Resistant calli can be seen ~8 weeks into selection.

7. Allow the resistant callus to grow before proceeding to the next step. A small amount can be saved as a back up in case contamination happens during regeneration.
    a. PREWORK: Make equivalent amounts of **R1-BAP** and **R1** containing 300mg/L Paramomycin. Pour onto 25mm plates and use a sharpie to divide the plate into sectors, this allows you to keep track of independent lines and grow multiple lines on each plate.

8. Move resistant callus onto **R1-BAP** containing 300mg/L Paramomycin. Incubate in dark for 3-4 days at 25C. After 3-4 days, place each callus onto **R1** containing 300mg/L Paramomycin. Incubate in the dark for 3-4 more days.

9. Regenerating callus may begin show signs of shoot and/or root growth. Remove regeneration plates into a growth chamber at 28C with 16h light and 8h dark. Cover plates with three layers of paper towels to create a low light environment. Each day, remove one paper towel until no paper towels remain. Check regularly for contamination. Roots and shoots will begin to develop during this time.
    a. PREWORK: Create **R2** media containing 150mg/L Paramomycin. Pour into magenta boxes or solo cups. Solo cups allow more vertical growth of the plants during the final tissue culture steps and let more light in.

10. 1-2 weeks after being in the light, **R1** cultures can be moved to **R2** containing 150mg/L Paramomycin. The cup or box can be divided so that 2-3 independent culture lines can fit. Allow these to grow until they either outgrow the cup or atleast 32 plants have a significant root and shoot system.
    a. PREWORK: Purchase market pack soil trays (08-04 at PSL storeroom), flats without holes, and 6 inch plastic domes. Fill market packs with soil used in greenhouse.

11. Move plants from **R2** into market packs with moist soil. Give each plant a unique identifier, usually denoted as "genotype: construct Event # - Plant # (e.g. ILP: Cas9_*ASNase* 32-3). Keep plants covered in dome to retain moisture and place in growth chamber. 3 days later, begin to move the dome so that ambient air can enter the flat. On the first day, move it so slightly that only the smallest amount of air can get in. The following days, adjust the dome more so that an increased amount of ambient air enters the flat. 4-5 days later the dome can be completely removed. Be sure to monitor the soil to ensure it does not dry out. 2-3 days later the plants can be moved into the greenhouse. Do not transplant them immediately; allow them to acclimate for 2-3 days before moving into 1000 size pots for T1 seed production.

APPENDIX C: Sequence Alignment for Plasmid and PCR Amplicons

Sequence alignment for plasmid gRNA assembly

>Reference *in silico* sequence
CTAGAAGTAGTCAAGGCGGCGAAGTATTCAGGCACGTGGCCAGGAAGAAGA
AAAGCCAAGACGACGAAAACAGGTAAGAGCTAAGCTTCCTGCAGGTTCACTG
CCGTATAGGCAGACAAGATGCTATGTATCAGCGTTTTAGAGCTAGAAATAGC
AAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCG
GTGCGTTCACTGCCGTATAGGCAGTGAGGAATCAAACATATGGTGTTTTAGA
GCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGT
GGCACCGAGTCGGTGCGTTCACTGCCGTATAGGCAGGAAAAATAACTGGTTA
CCCGGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCA
ACTTGAAAAAGTGGCACCGAGTCGGTGCGTTCACTGCCGTATAGGCAGGGAG
GCCTACGGGTTCTGCTGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCT
AGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGTTCACTGCCGT
ATAGGCAGGTCGATCGACAAGCTCGAGTTTCTCCATAATAATGTGTGAGTAG
TTCCCAGATAAGGGAATTAGGGTTCC

>Lw1 Clone 7 TC089 F primer
CTAGAAGTAGTCAAGGCGGCGAAGTATTCAGGCACGTGGCCAGGAAGAAGA
AAAGCCAAGACGACGAAAACAGGTAAGAGCTAAGCTTCCTGCAGGTTCACTG
CCGTATAGGCAGACAAGATGCTATGTATCAGCGTTTTAGAGCTAGAAATAGC
AAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCG
GTGCGTTCACTGCCGTATAGGCAGTGAGGAATCAAACATATGGTGTTTTAGA
GCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGT
GGCACCGAGTCGGTGCGTTCACTGCCGTATAGGCAGGAAAAATAACTGGTTA
CCCGGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCA
ACTTGAAAAAGTGGCACCGAGTCGGTGCGTTCACTGCCGTATAGGCAGGGAG
GCCTACGGGTTCTGCTGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCT
AGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGTTCACTGCCGT
ATAGGCAGGTCGATC

>Lw1 Clone 7 TC320 R primer
GAAGAGAAAGCCAGACGACGAAACAGGTAAGAGCTAAGCTTCCTGCAGGTT
CACTGCCGTATAGGCAGACAAGATGCTATGTATCAGCGTTTTAGAGCTAGAA
ATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCG
AGTCGGTGCGTTCACTGCCGTATAGGCAGTGAGGAATCAAACATATGGTGTT
TTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTATCAACTTGAA
AAAGTGGCACCGAGTCGGTGCGTTCACTGCCGTATAGGCAGGAAAAATAACT
GGTTACCCGGTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGT
TATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGTTCACTGCCGTATAGGCA
GGGAGGCCTACGGGTTCTGCTGTTTTAGAGCTAGAAATAGCAAGTTAAAATA
AGGCTAGTCCGTTATCAACTTGAAAAAGTGGCACCGAGTCGGTGCGTTCACT
GCCGTATAGGCAGGTCGATCGACAAGCTCGAGTTTCTCCATAATAATGTGTG
AGTAGTTCCCAGATAAGGGAATTAGGGTTCC

Sequence alignment for *Lemon White 1* and *Asparaginase* plants

>Wildtype sequence Lw1-F1 – Lw1-R2
CACTTCATGGCCTTCAATACTATTTGTGATGCCACTCAGGTACTTTTTCCTTTT
TCCAACATTTCCTTCAGCTTTGCATAATACAAAGTCTTGTAGCCCAACCCAAC
CGAGTCTCTTTCTTTTAGATGTTACGTGAAGCATTCCAAAATGTATGGTTTCTT
GATGATTTAAAAACTGGGAATGATGGATTTAGCACTACTGTACTGTAACGTCT
AGAAAAGGCTAGCGTACAATTCATGATAGACGGTAGCCATTCATGGTATAAT
ACAATCTCCATCACCCTTTACATGCAACCTGCAAAGGAAACTGAGGCAACTA
GATGGCACTTTTATTTGGGATCAAAGTTAAGAGGATCCTAGTCCATTTGTTGC
TCCACTTTTGGTTTGAGAAAATGACTCTTTGGGTGATAGATTTATAATTGTTTT
GACATATGCTGCATTGACTTATATACATCCGACCATATGTTTGATTCCTCAGG
AAAGACAAGATGCTATGTATCAGCTGGTGAAAGAGAAAGTTGACCTTATTCT
TGTTGTTGGAGGATGGAATTCAAGTAACACCTCTCATCTGCAAGAAATCGGA
GAACTCAGTGGAATTCCATCATACTGGATTGACAGTGAACAAAGGATTGGAC
CAGGAAACAGGATCAGCTACAAGTTAAATGTAATTTCTCATACCTTCCACAG
CATTCATTTTTTAAAAATAATAGAATTTTGAAGGGACCTTCTTTTAATTAGAA
ACTGGAAAGAGCTTAAATTTAATTATTGAAAGTATATGATCTCATTTTAACCG
AAAATTGGATTGTTTAACAGCATGGTGAACTGGTTGAGAAAAATAACTGGTT
ACCCGAGGGGCCTATTACCATTGGTGTTACTTCAGGTGCCTCAACTCCAGATA
AGGT

>ILP E1 (800bp band) 3’
TTTTTNTTTTTTTTCCCAACATTTTCCCTTTCAGGCTTTTGCATAATANCAAAG
TTCTTGTAGNCCCNACCCCAACCGAGTCTCTTTTTTTTTAGATGTTACGTGAA
GCATTNCCAAAAATGTATGGGTTTCTTGATGATTTTAAAAAANCTGGGAATG
ATGGATTTTAGCACTACTGTACTGTAACGTCTAGAAAAGGCTAGCGTACAATT
CATGATAGACGGTAGCCATTCATGGTATAATACAATCTCCATCACCCTTTACA
TGCAACCTGCAAAGGAAACTGAGGCAACTAGATGGCACTTTTATTTGGGATC
AAAGTTAAGAGGATCCTAGTCCATTTGTTGCTCCACTTTTGGTTTGAGAAAAT
GACTCTTTGGGTGATAGATTTATAATTGTTTTGACATATGCTGCATTGACTTAT
ATACATCCGACCATGTTTGATTCCTCAGGAAAGACAAGATGCTATGTACTGGT
GAAAGAGAAAGTTGACCTTATTCTTGTTGTTGGAGGATGGAATTCAAGCAAC
ACCTCTCATCTGCAAGAAATTGGAGAACTCAGTGGAATTCCATCATACTGGA
TTGACAGTGAACAAAGGATTGGACCAGGAAACAGGATCAGCTACAAGCTAA
ATGTAATTTCTCATACCTTCCACAGCATTCATTTTTTGAAAAATAATAGAATT
TTGAAGGGACATTCCTTTTAATTAGAAACTGGAAAGAGCTTGAATTTATTTAT
TGAAAGTATATGATCTCATTTTTAACCGAAAATTGGATTGTTTAACAGCATGG
TGAACTGGCCTATTACCATTGGTGTTACTTCAGGTGCCTCAACTCCAGATAAG
GT

>ILP E1 (500bp band) 3’
TTTCTTTTTCCACATTTCCTTCAGCTTGCATAATACAAAGTCTTGTAGCCCAAC
CCAACCGAGTCTCTTTCTTTTAGATGTTACGTGAAGCATTCCAAAATGTATGG
TTTCTTGATGATTTAAAAAACTGGGAATGATGGATTTAGCACTACTGTACTGT
AACGTCTAGAAAAGGCTAGCGTACAATTCATGATAGACGGTAGCCATTCATG

GTATAATACAATCTCCATCACCCTTTACATGCAACCTGCAAAGGAAACTGAG
GCAACTAGATGGCACTTTTATTTGGGATCAAAGTTAAGAGGATCCTAGTCCAT
TTGTTGCTCCACTTTTGGTTTGAGAAAATGACTCTTTGGGTGATAGATTTATA
ATTGTTTTGACATATGCTGCATTGACTTATATACATCCGACCACCCGAGGGGC
CTATTACCATTGGTGTTACTTCAGGTGCCTCAACTCNNNATAAGGT

>ILP E2 3’
GGCCTNCAATACTATTTTGGNGATGNCCNCCTCCAGGGTAATTTTTTTCTTTTT
TTCCNACCATTTCCNTTCAGCTTNGCATTAATACCAAAGTTCTTGGTAGGCCC
AACCCNANCCGAGTTTCTTTCTTTTNNGATGTTACGTGAAGCATTCCAAAATG
TATGGTTTCTTGATGATTTAAAAAACTGGGAATGATGGATTTAGCACTACTGT
ACTGTAACGTCTAGAAAAGGCTAGCGTACAATTCNTGATAGACGGTAGCCAT
TCATGGTATAATACAATCTCCATCACCCTTTACATGCAACCTGCAAAGGAAAC
TGAGGCAACTAGATGGCACTTTTATTTGGGATCAAAGTTAAGAGGATCCTAG
TCCATTTGTTGCTCCACTTTTGGTTTGAGAAAATGACTCTTTGGGTGATAGATT
TATAATTGTTTTGACATATGCTGCATTGACTTATATACATCCGACCAATATGT
TTGATTCCTCAGGAAAGACAAGATACATGTTTGATTCCTCAGGAAAGACAAG
ATGCTATGTAAGCTGGTGAAAGAGAAAGTTGACCTTATTCTTGTTGTTGGAGG
ATGGAATTCAAGTAACACCTCTCATCTGCAAGAAATTGGAGAACTCAGTGGA
ATTCCATCATACTGGATTGACAGTGAACAAAGGATTGGACCAGGAAACAGGA
TCAGCTACAAGCTAAATGTAATTTCTCATACCTTCCACAGCATTCATTTTTTG
AAAAATAATAGAATTTTGAANGGGACATTCCTTTTAATTAGAAACTGGAAAG
AGCTTAAATTTATTTATTGAAANGTATATGATCTCATTTTTAANCCGAAAATT
GGATTGTTTAACCCGCATGGTGAANCTGGTTGAGAAAAATAANCTGGTANCC
CGGGGGCCT

>H99 E1 3’
TNTCTTTTCNTTTTANGATGTTACGTGNAAGNNNTCCCAAAATGTATGGNTTT
TCTTGATGATTTTAAAAAAAACTGGGAATGATGGATTTAGCACTACTGTACTG
TAACGTCTAGAAAAAGGCTAGCGTACAATTCCATGATAGACGGTAGCCATTC
ATGGTATAATACAATCTCCATCACCCTTTACATGCAACCTGCAAAGGAAACT
GAGGCAACTAGATGGCACTTTTATTTGGGATCAAAGTTAAGAGGATCCTAGT
CCATTTGTTGCTCCACTTTTGGTTTGAGAAAATGACTCTTTGGGTGATAGATTT
ATAATTGTTTTGACATATGCTGCATTGACTTATATACATCCGACCAGCTGGCG
AAAGAGAAAGTTGACCTTATTCTTGTTGTTGGAGGATGGAATTCAAGTAACA
CCTCTCATCTGCAAGAAATTGGAGAACTCAGTGGAATTCCATCATACTGGATT
GACAGTGAACAAAGGATTGGACCAGGAAACAGGATCAGCTACAAGCTAAAT
GTAATTTCTCATACCTTCCACAGCATTCATTTTTTGAAAAATAATAGAATTTT
GAAGGGACATTCCTTTTAATTAGAAACTGGAAAGAGCTTAAATTTATTTATTG
AAAGTATATGATCTCATTTTTAACCGAAAATTGGATTGTTTAACAGCATGGTG
TACTGGTTGAGAAAAATAAACTGGTTACCGAGGGGCCTATTACCATTGGTGT
TACTTCAGGTGCCTCAACTCCAGATAN

> Wildtype Sequence Lw1-F2 Lw1- R5
GCAACCAGTACACCTGGGGCCCCGTCACCGTGAAGCTCGCGGAGGCCTACGG
GTTCTGCTGGGGCGTCGAGCGCGCCGTGCAGATCGCGTACGAGGCGCGCAAG

CAGTTCCCCGAGGAGCGCATCTGGCTCACCAACGAAATCATCCACAACCCCA
CCGTCAACAAGGTAACCGTATCCGAAATCTGTGAACTCTGACCAAGAATTTG
AACTTCGTATGGAACTAGGTGGGTGACGAAATTAGTGGTCTGCTGCATCCTCT
TGTGTATATCAGAGAGCAGAACGATCTGTTGACCCACTAAGCGTTACAAAGT
AACA

>ILP E1 5'
NCNACCCNTNCNCCNNNGNCNCCGTCANCNTNNNGNNNGCGNANGNGTTCG
GNTNGGGCGTCGAGCGCGCCGTGCAGATCGCGTACGAGGCGCGCAAGCAGT
TCCCCGAGGAGCGCATCTGGCTCACCAACGAAATCATCCACAACCCCACCGT
CAACAAGGTAACCTGTGAACTCTGACCAAGAATTTGAGCTTCGTATGGAACT
AGGTGGGTGAGGAAATCAGTGGTCTGCTGCATCCTCTTGTGTATATCAAGAG
AGCAGAACGAGCTGTTGACCCACTAAGGCGGTACAAAGTAACA

>ILP E2 5'
GCAACCAGTACACCTGGGGCCCCGTCACCGTGAAGCTCGCGGAGGCCTACGG
GTGGGGCGTCGAGCGCGCCGTGCAGATCGCGTACGAGGCGCGCAAGCAGTTC
CCCGAGGAGCGCATCTGGCTCACCAACGAAATCATCCACAACCCCACCGTCA
ACAAGGTAACCTGTGAACTCTGACCAAGAATTTGAGCTTCGTATGGAACTAG
GTGGGTGAGGAAATCAGTGGTCTGCTGCATCCTCTTGTGTATATCAAGAGAG
CAGAACGAGCTGTTGACCCACTAAGCGGTACAAAGTAACA

>H99 E1 5'
GNNANCCNNNCCCCNNGGGNCCCNTNCCCNNNNANNTTNNNGNNGNCTTCN
GGNTTNNGCTGGGGCGTCGAGCGCGCCGTGCAGATCGCGTACGAGGCGCGC
AAGCAGTTCCCCGAGGAGCGCATCTGGCTCACCAACGAAATCATCCACAACC
CCACCGTCAACAAGGTAACCTGTGAACTCTGACCAAGAATTTGAGCTTCGTA
TGGAACTAGGTGGGTGAGGAAATCAGTGGTCTGCTGCATCCTCTTGTGTATAT
CAAGAGAGCAGAACGAGCTGTTGACCCACTAAGCGGTACAAAGTAACA

>ILP WT *ASNase*
CTGATGTTGGTTCGCTATTTTTCTCTAGGGCCTGGAGACCGTGGACAACAGCT
ACTTCATCACGGANGACAACGTCGGCATGCTCAAGCTCGCCAAGGAGGCCGG
CAGCATCCTGGTTGGTTCCGTCCGCCTACAACACCCACTCCTGTCAACTTTTT
TTCTTTTCCTTTTGGGTTTCCGAAATCAATCCAGAGGTGGTGACCTGACTGAC
CACGCACGGATGTGGGCGCGCGTGCAGTTCGACTACCGGATCCCGCTGGCNG
GGACGGACACGTGCAGCGCGCTGGCGGGCGCGGCGGACAGCAACGGCGGCG
GCGTGCACAAGGCGGGGATGGTCATGAACGGGCTGCCCATCAGCGTGTACGC
GCCGGAGACGGTGGGGTGCGCGGTGGTGGACGCGNCGGGGGCCTGCGCGGC
GGCCACGTCCACGGGCGGGCTCATGAACAAGATGGCCGGCCGCATCGGGGA
CTCGCCGCTCATCGGCTCCGGCACCTACGCGTGCGGCGCCTGCGCCGTGTCGT
GCACGGGCGAGGGCGAGGCCATCATCCGCTNCACGCTCGCGCGCGACGTGGC
CGCCGTCATGGAGTACAAGGGNCTGCCCCTGCAGNAGGCCGTCGACTACTGC
GTCAAGGAGCGCCTCGACCAGGGATTCGCCGGCCTCATCGCCGTCTCCGGCA
CCGGCGAGGTGGCGTACGGATTCAACTGCACCGGCATGTTCCGGGGCTGCGC
CACCGAGGACGGGTTCATGGAGGTCGGCATCTGGGAGTGAGCTGGAGGCCCC

GCCATCTAAACACAAATCAAAATCAGTGTGCGTGCAGCTGCATGTGGCCATC
CTGCGATTGCCATGTGTTTAATCTCCGGGTCACAGGTGTTGTCGATGA

>ILP ASN 1
TTTTTTCTTTAGGGCCTNGAGACGCTNGCAGGGACCGGACACGTGCAGCCCG
CCTGGCGGGCGCGGCGGACAGCAACGGCGGCGGCGTGCNCAAGGCGGGGAT
GGTCATGAACGGGCTGCCCATCAGCGTGTACGCGCCNGAGACGGTGGGGTGC
GCGGTGGTGGACGCGACGGGGGCCTGCGCGGCGGCCACGTCCACGGGCGGG
CTCATGAACAAGTTGGCCGGCCGCATCGGGGACTCGCCGCTCATCGGCTCCG
GCACCTACGCGTGCGGCGCCTGCGCCGTGTCGTGCACGGGCGAGGGCGAGGC
CATCATCCGCTCCACGCTCACNNGCGACGTGGCCGCCGCCATGGAGTACAAG
GGCCTGCCCCTGCAGGAGGCCGTCGACTACTGCGTCAAGGAGCGCCTCGACC
AGGGATTCGCCGGCCTCATCGCCGTCTCCGGCACCGGCGTGGCGTACGGATT
CAACTGCACCGGCATGTTCCGGGGCTGCGCCACCGAGGACGGGTTCATGGAG
GTCGGCATCTGGGAGTGAGCAGCGCGGCCGGGGCGGAGCTGGAGGCCCGCC
ATCTAAAAGACAAATCAAAATNCAGTGTGCGTGCAGCTGCATGTGGCCATCC
TGCGNTTNAACNGTGTGTTTAATCTCCGGGTCACAGGTGTTGTC

>ILP ASN 2
CTGATGTTGATTCGCTATTTTTCTCTAGGGCCTGGAGACACTGGCAGGGACGG
ACACGTGCAGCGCGCTGGCGGGCGCGGCGGACAGCAACGGCGGCGGCGTGC
ACAAGGCGGGGATGGTCATGAACGGGCTGCCCATCAGCGTGTACGCGCCGG
AGACGGTGGGGTGCGCGGTGGTGGACGCGACGGGGGCCTGCGCGGCGGCCA
CGTCCACGGGCGGGCTCATGAACAAGATGGCCGGCCGCATCGGGGACTCGCC
GCTCATCGGCTCCGGCACCTACGCGTGCGGCGCCTGCGCCGTGTCGTGCACG
GGCGAGGGCGAGGCCATCATCCGCTCCACGCTCGCGCGCGACGTGGCCGCCG
TCATGGAGTACAAGGGCCTGCCCCTGCAGGAAGGCCGTCGACTACTGCGTCA
AGGAAGCGCCTCGACCAGGGATTCGCCGGCCTCATCGCCGTCTCCGGCGCCG
GCGTGGCGTACGGATTCAAACTGCACCGGCATGTTCCGGGGGCTGCGCCACC
GAGGAACGGGTTCATGGAAGGTCGGCATCTGGGA